https://doi.org/10.31891/csit-2025-1-5
UDC 004.89

Serhii HLADIHOLOV, Oleksii KOZACHKO
Vinnytsia National Technical University

PREDICTION OF ALZHEIMER'S DISEASE USING BAYESIAN NEURAL NETWORKS

This article presents a methodology for optimizing Bayesian neural networks and their application to complex prediction tasks, with a focus on diagnosing Alzheimer's disease. Alzheimer's is a neurodegenerative condition where early detection is vital for initiating timely interventions and improving patient outcomes. The proposed methodology includes determining the optimal structure of classical neural networks by performing grid search to identify the best combination of layers and neurons. The architecture identified through cross-validation forms the basis for constructing Bayesian neural networks, where weights derived from classical models are utilized as prior distributions. This integration improves prediction accuracy while preserving the Bayesian network's capacity for quantifying uncertainty.

Bayesian models are trained using Markov Chain Monte Carlo methods, with experiments exploring the impact of prior distribution parameters, including variations in means and standard deviations. Results show that a mean value of zero and a standard deviation of 2.5 yield optimal outcomes, minimizing classification error while balancing uncertainty estimation. Increasing the standard deviation improved performance up to a threshold, beyond which further gains were statistically insignificant. The ability of Bayesian neural networks to incorporate uncertainty provides critical advantages for decision-making in medical contexts, particularly in scenarios involving incomplete or noisy data.

The findings demonstrate that Bayesian neural networks based on optimized classical architectures can effectively address prediction tasks in high-stakes domains like medicine. By leveraging prior knowledge, the proposed approach reduces training time and enhances model performance, offering a robust framework for diagnosing Alzheimer's disease. Future research will explore automating structural optimization, assessing the impact of different prior distributions, and extending this methodology to other neurodegenerative disorders.

Keywords: Bayesian neural networks, Alzheimer's disease prediction, machine learning, a priori distributions, deep learning.

Сергій ГЛАДІГОЛОВ, Олексій КОЗАЧКО
Вінницький національний технічний університет

ПЕРЕДБАЧЕННЯ ХВОРОБИ АЛЬЦГЕЙМЕРА ЗА ДОПОМОГОЮ БАЙЄСОВИХ НЕЙРОННИХ МЕРЕЖ

У цій статті представлено методологію оптимізації байєсівських нейронних мереж та їх застосування до комплексних задач прогнозування, з акцентом на діагностиці хвороби Альцгеймера. Хвороба Альцгеймера є нейродегенеративним станом, раннє виявлення якого є життєво важливим для початку своєчасного втручання та покращення результатів лікування. Запропонована методологія включає визначення оптимальної структури класичних нейронних мереж шляхом проведення пошуку по сітці для виявлення найкращої комбінації шарів і нейронів. Архітектура, визначена шляхом крос-валідації, є основою для побудови байєсівських нейронних мереж, де ваги, отримані з класичних моделей, використовуються в якості параметрів апріорних розподілів. Така інтеграція підвищує точність прогнозування, зберігаючи при цьому здатність байєсівської мережі до кількісної оцінки невизначеності.

Байєсівські моделі навчаються за допомогою методу Монте-Карло марковських ланцюгів, а експерименти досліджують вплив апріорних параметрів розподілу, включаючи варіації середніх значень і стандартних відхилень. Результати показують, що середнє значення, рівне нулю, і стандартне відхилення 2.5 дають оптимальні результати, мінімізуючи похибку класифікації і збалансовуючи оцінку невизначеності. Збільшення стандартного відхилення покращує продуктивність до певного порогового значення, за яким подальший приріст є статистично незначущим. Здатність байєсівських нейронних мереж враховувати невизначеність надає критичні переваги для прийняття рішень у медичному контексті, особливо в сценаріях, пов'язаних з неповними або зашумленими даними.

Результати дослідження демонструють, що байєсівські нейронні мережі на основі оптимізованих класичних архітектур можуть ефективно вирішувати задачі прогнозування в таких відповідальних сферах, як медицина. Використовуючи апріорні знання, запропонований підхід скорочує час навчання і підвищує продуктивність моделі, пропонуючи надійну основу для діагностики хвороби Альцгеймера. Подальші дослідження будуть присвячені автоматизації структурної оптимізації, оцінці впливу різних апріорних розподілів та поширенню цієї методології на інші нейродегенеративні розлади.

Ключові слова: Байєсові нейронні мережі, прогнозування хвороби Альцгеймера, машинне навчання, апріорні розподіли, глибоке навчання.

Introduction

Alzheimer's disease is one of the most widespread neurodegenerative conditions worldwide, characterized by severe cognitive decline that eventually leads to a complete loss of independence. Since the disease progresses irreversibly, current treatment strategies are limited to slowing its advancement. As such, early detection is crucial to enable timely therapeutic measures, which can substantially improve the quality of life for those affected [1]. Given the disease's complexity and the multifactorial nature of its underlying causes, traditional predictive models like logistic regression often fail to capture all interdependencies and risk factors effectively. In contrast, neural networks, which have already shown significant success in addressing classification and prediction challenges

across various domains, offer a promising alternative for the early diagnosis of Alzheimer's disease. However, these models do not provide uncertainty estimates, which is a crucial aspect in medical diagnostics [2].

At the same time, Bayesian neural networks, which combine the power of neural networks with the probabilistic approach of Bayesian methods, allow not only for making predictions but also for quantitatively assessing the confidence in the results. This enables the consideration of risks and facilitates more informed decision-making in uncertain conditions, which is particularly important in the medical field. Thus, the task lies in exploring the potential of Bayesian neural networks for predicting Alzheimer's disease and comparing their effectiveness with classical methods [3].

Analysis of Research and Publications

The early diagnosis of Alzheimer's disease remains a prominent focus in both medical diagnostics and machine learning research. Numerous studies emphasize the application of diverse models to facilitate early detection of the disease, including both classical statistical approaches and modern deep learning methods [4].

Classical methods, such as logistic regression, have long been used to model the probability of disease occurrence based on a set of predictors, including demographic, clinical, and genetic factors. For example, in studies [5, 6, 7], the effectiveness of logistic regression was demonstrated for predicting Alzheimer's disease using traditional clinical data, such as the results of neuropsychological tests. However, these methods are often limited in their ability to account for complex nonlinear interactions between features.

In light of this, an increasing number of researchers are turning to neural networks, which can model complex nonlinear relationships. Studies [8, 9, 10, 11] have shown that classical neural networks can achieve high accuracy in predicting Alzheimer's disease using both structured clinical data and unstructured data, such as medical images. However, the main issue with neural networks is the lack of tools for providing uncertainty estimates in their predictions, limiting their use in critical areas such as medicine.

Bayesian neural networks offer a solution to this problem by integrating the principles of the Bayesian probabilistic approach into deep learning. They not only allow for making predictions but also provide information about the uncertainty of those predictions. Studies [12, 13, 14] have shown that Bayesian neural networks can be effective in medical diagnostics, allowing uncertainty in predictions to be accounted for, and supporting more informed decision-making in high-risk situations.

However, as studies [15, 16] indicate, Bayesian neural networks often underperform compared to classical neural networks in terms of prediction accuracy due to the complexity of optimization and the selection of prior distributions. Special attention is given to the use of prior information from classical neural networks for initializing Bayesian models [17]. This approach allows for improved prediction accuracy while retaining the advantages of uncertainty estimation, significantly enhancing the effectiveness of Bayesian neural networks in medical diagnostic tasks, particularly in predicting Alzheimer's disease.

Formulation of the Article's Objectives

The goal of this work is to develop a methodology for constructing and optimizing the structure of Bayesian neural networks for predicting Alzheimer's disease, taking into account the longer training time compared to classical neural networks. The research aims to:

1) Determine the optimal architecture of a classical neural network using grid search, varying the number of layers and neurons to obtain the most efficient structure.
2) Compare the performance of models based on the F1 metric using cross-validation to identify the best classical neural network architecture.
3) Develop a Bayesian neural network using the optimal architecture derived from the classical network.
4) Perform experiments with various prior distributions in Bayesian neural networks to analyse their effects on prediction accuracy and uncertainty estimation.
5) Identify the most suitable parameters for prior distributions that reduce classification errors while maintaining high prediction accuracy.

This approach aims to enhance not only the overall effectiveness of Bayesian neural networks but also reduce their training time, which is critically important for their further application in medical diagnostics.

Presentation of the Main Material

For the research, the "Alzheimer's Disease Dataset" [18], available on the Kaggle platform, was used. This dataset contains information on six categories of patients. These categories include demographic data such as gender, age, education, and ethnicity; lifestyle-related information such as body mass index, smoking, alcohol consumption, hours of physical activity, and the quality of diet and sleep; the presence of other diseases in the patient or family history; clinical measurements of blood pressure and lab tests; and cognitive ability assessments. The target variable is the diagnosed presence of Alzheimer's disease. The dataset consists of 2,149 records.

To ensure high-quality predictive models and avoid potential errors during the training of machine learning algorithms, data preprocessing was performed at the initial stage. Below are the key steps of this process:

1. Removing technical information: The dataset contained two columns with anonymized patient data. Since these columns did not provide useful information for further analysis, they were removed.

2. Data scaling**:** As the features in the dataset had different ranges of values, scaling methods were applied to normalize the data using the MinMaxScaler. This approach transforms the feature values into the range of 0 to 1, ensuring an even contribution of each feature to the modelling and facilitating faster and more stable performance of the models.

3. Class balancing: Since the original dataset had an imbalance between classes (patients with Alzheimer's disease and those without it, see Fig. 1), the oversampling method was applied to balance the classes by increasing the number of examples in the underrepresented class.

These steps helped reduce the risk of overfitting to the majority class and ensured more accurate predictions.

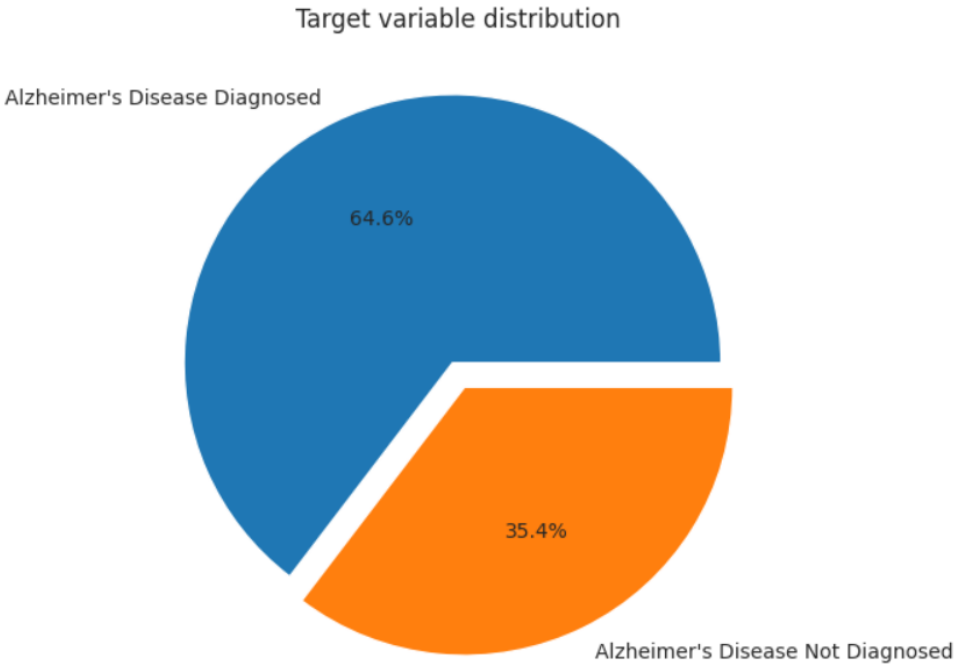

Fig. 1. Target variable class distribution

4. Splitting into train and test sets**:** The data was split into training and test sets in an 80:20 ratio. Training was conducted using the cross-validation technique.

Let's formalize the binary classification task for predicting Alzheimer's disease. Let us define

$$D = \{(x_i, y_i)\}^N_{i=0}, \quad (1)$$

where:

$x_i \in R^d$ is the feature vector of patient *i* from the space R^d;

$y_i \in \{0, 1\}$ is the class label (1 - Alzheimer's diagnosed, 0 - not diagnosed) for patient *i*;

N is the number of examples in the dataset (2149);

d is the number of features after data preparation (32).

For the neural network model, we need to find $f(x, \theta): R^d \to [0, 1]$, where θ represents the model parameters predicting the membership of vector x_i to class 1.

The function $f(x, \theta)$ consists of L layers, and for each layer l, the following are defined:

$z^{(l)} \in R^{n_l}$ is the vector of linear combinations in layer *l*, where n_l is the number of neurons in the layer;

$a^{(l)} \in R^{n_l}$ is the activation of neurons after applying the activation function to $z^{(l)}$, $a^{(0)} = x$;

$W^{(l)} \in R^{n_l \times n_{l-1}}$ is the weight matrix for connections between layers *l* and *(l-1)*;

$b^{(l)} \in R^{n_l}$ is the bias vector for layer *l*.

For layer *l*, the output linear combination for the neurons is computed as:

$$z^{(l)} = W^{(l)} a^{(l-1)} + b^{(l)} \quad (2)$$

The activation of each layer is computed as:

$$a^{(l)} = \sigma\left(z^{(l)}\right), \tag{3}$$

where $\sigma(z)$ is the sigmoid activation function.

The error function for binary classification is defined using cross-entropy as:

$$L(\theta) = -\frac{1}{N}\sum_{i=0}^{N}\left(y_i\ log\ (\hat{y}_i) + (1-y_i)\ log\ (1-\hat{y}_i)\right), \tag{4}$$

where $\hat{y}_i = a^{(L)}$ is the predicted class by the model, and y_i is the true class label.

Model training is performed using the gradient descent method:

$$\theta^{(t+1)} = \theta^{(t)} - \eta\nabla_\theta L(\theta), \tag{5}$$

where η is the learning rate coefficient, and $\nabla_\theta L(\theta)$ is the gradient of the error function with respect to the parameters θ.

Bayesian neural networks are based on the Bayesian approach to probability theory and Bayes' theorem:

$$p(\theta|D) = \frac{p(\theta)p(D|\theta)}{p(D)}, \tag{6}$$

where $p(\theta)$ is the prior distribution of the parameters, $p(\theta|D)$ is the posterior distribution of the parameters, $p(D|\theta)$ is the likelihood, and $p(D)$ is the normalizing constant, representing the probability of the data integrated over all possible parameter values.

We define the prior distributions of the Bayesian neural network parameters $\theta = \{W^{(l)}, b^{(l)}\}$, which reflect the initial belief about the parameter values:

$$W_{jk}{}^{(l)} \sim N(\mu_w, {\sigma_W}^2), b_j{}^{(l)} \sim N(\mu_b, {\sigma_b}^2), \tag{7}$$

where μ_w, μ_b are the mean values, and σ_W, σ_b are the standard deviations.

For binary classification, the distribution of the target variable y is defined as a Bernoulli distribution with parameter p, equal to the output of the last layer of the network:

$$y \sim Ber\left(\sigma\left(W^L a^{(L-1)}\right)\right) \tag{8}$$

Training of Bayesian models is performed using Markov Chain Monte Carlo (MCMC) methods.

The practical experiments followed the structure below:

1) Selection of neural network architectures: To determine the optimal architecture of the classical neural network, the grid search method was applied. During the search, the number of layers and the number of neurons in each layer were varied to identify the most efficient combination of parameters.

2) Determining the optimal model: Cross-validation was used to assess the efficiency of the models, which helped identify the best neural network architecture based on the averaged F1 score metrics.

3) Construction of the Bayesian neural network: Based on the identified architecture of the classical neural network, an equivalent Bayesian neural network was constructed.

4) Testing with different prior distributions: A grid search of the prior distribution parameters for the Bayesian neural network was conducted, where the mean values and standard deviations were varied.

5) Selecting the best parameters: Based on the test results, the optimal prior parameters were selected, which minimized classification error and maximized prediction accuracy.

The models were implemented using the Python programming language and the libraries Scikit-learn and PyMC. The scheme of the constructed Bayesian neural network is shown in Fig. 2.

The experimental results indicated that the best architecture for the classical neural network was a two-layer structure with 10 neurons in each layer.

Experiments with the parameter search for the Bayesian neural network showed that with increasing uncertainty, due to a higher standard deviation of the prior distribution, the training time of the model increased, while its performance improved. However, when the deviation exceeded 2.5, the improvement in prediction accuracy was within the statistical margin of error. The optimal prior distribution parameters were found to be a mean value of 0 and a standard deviation of 2.5. The optimal mean of 0 can be explained by the fact that both positive and negative weights are equally important for the model, and deviations from 0 lead to a decrease in model quality. The performance metrics of the best neural network, as well as the best and worst Bayesian neural networks, are presented in Table 1.

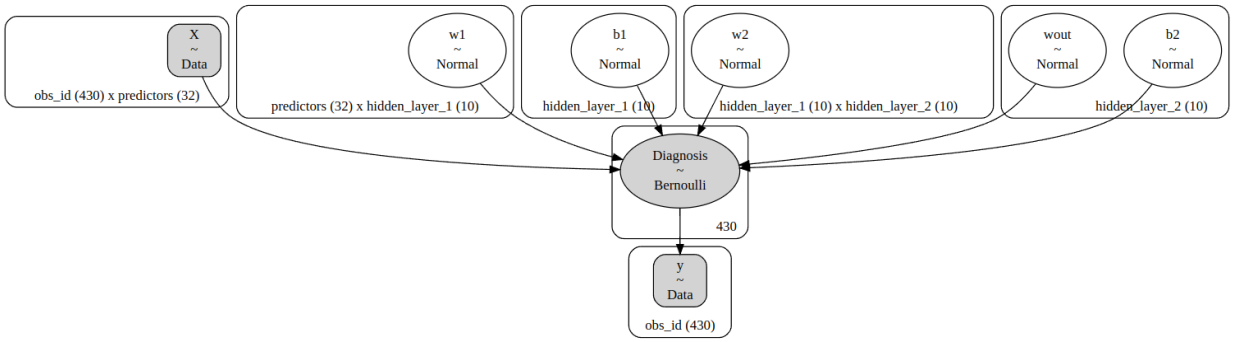

Fig. 2. Bayesian neural network architecture scheme

Table 1

Model performance metrics

	Accuracy	**Precision**	**Recall**	**F1**
Classical Neural Network	0.8209	0.7354	0.7535	0.7181
Bayesian Neural Network ($\mu = 1, \sigma = 0.5$)	0.8047	0.9158	0.5337	0.6744
Bayesian Neural Network ($\mu = 0, \sigma = 2.5$)	**0.9000**	**0.9225**	**0.8037**	**0.8590**

Conclusions from this research and prospects for further studies

In the course of this research, a methodology for searching the structure of Bayesian neural networks was developed, taking into account the training duration of these models in comparison to classical neural networks. In particular, it was demonstrated that initially determining the optimal number of layers and neurons in classical neural networks is an effective approach to reducing training time and improving the performance of Bayesian models.

The beginning of the work was focused on finding the optimal architecture of the classical neural network, which provided the necessary parameters for further work with Bayesian networks. This stage included a detailed analysis of different architectures and key parameters which include the number of layers, the quantity of neurons per layer, and the implementation of regularization techniques to mitigate overfitting.

Future directions for research in this domain:

1. **Enhancing methodological approaches:** Creating novel algorithms to automate the process of identifying optimal Bayesian neural network architectures could substantially decrease training time and enhance model performance.
2. **Examining prior distribution effects:** Analysing the impact of diverse prior distribution types on the efficacy of Bayesian neural networks, particularly in the context of medical data, may offer new opportunities for improving models.
3. **Combined approaches:** Integrating results obtained using other machine learning algorithms, such as decision trees or boosting methods, to improve the estimation of distribution parameters could become a promising research direction.
4. **Extended application in medical practice:** Exploring the possibilities of applying this methodology to other neurodegenerative diseases, as well as to clinical practice in general, could significantly expand the application boundaries of the developed models.
5. **Uncertainty estimation in predictions:** Further research on methods for estimating the uncertainty of predictions made by Bayesian neural networks could assist doctors in making more informed decisions in complex clinical situations.

References

1. Nandi, A., Counts, N. Z., Chen, S., Seligman, B., Tortorice, D., Vigo, D., & Bloom, D. E. (2022). Global and regional projections of the economic burden of Alzheimer's disease and related dementias from 2019 to 2050: A value of statistical life approach. *eClinicalMedicine, 51*. URL: https://api.semanticscholar.org/CorpusID:250996264
2. Yang, K., & Mohammed, E. (December 2020). A Review of Artificial Intelligence Technologies for Early Prediction of Alzheimer's Disease. *A Review of Artificial Intelligence Technologies for Early Prediction of Alzheimer's Disease*. doi:10.48550/arXiv.2101.01781
3. Pérez-Millan, A., Contador, J., Tudela, R., Niñerola-Baizán, A., Setoain, X., Lladó, A., . . . Sala-Llonch, R. (August 2022 p.). Evaluating the performance of Bayesian and frequentist approaches for longitudinal modeling: application to Alzheimer's disease. *Scientific Reports, 12*. doi:10.1038/s41598-022-18129-4
4. Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., & Lin, C. T. (April 2020 p.). Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review. *ACM Trans. Multimedia Comput. Commun. Appl., 16*. doi:10.1145/3344998
5. Memon, M. H., Li, J., Haq, A. U., & Hunain Memon, M. (2019). Early Stage Alzheimer's Disease Diagnosis Method. *2019 16th International Computer Conference on Wavelet Active Media Technology and Information Processing*, (cc. 222-225). doi:10.1109/ICCWAMTIP47768.2019.9067689

6. R.Sendhil, D., Patibandha, N., Mandal, A., & Shah, P. (2025). Comparative Evaluation of Machine Learning Techniques for Early Alzheimer's Detection: Challenges and Insights. *2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI)*, 1050-1058. URL: https://api.semanticscholar.org/CorpusID:276675785
7. Lazarova, S., Grigorova, D., & Petrova-Antonova, D. (2023). Detection of Alzheimer's Disease Using Logistic Regression and Clock Drawing Errors. *Brain Sciences, 13*. URL: https://api.semanticscholar.org/CorpusID:260360996
8. Bi, X., Liu, W., Liu, H., & Shang, Q. (2021). Artificial Intelligence-based MRI Images for Brain in Prediction of Alzheimer′ s Disease. *Journal of Healthcare Engineering, 2021*, 8198552.
9. Khojaste-Sarakhsi, M., Haghighi, S. S., Ghomi, S. F., & Marchiori, E. (2022). Deep learning for Alzheimer's disease diagnosis: A survey. *Artificial intelligence in medicine, 130*, 102332.
10. Marzban, E. N., Eldeib, A. M., Yassine, I. A., Kadah, Y. M., & Initiative, A. D. (2020). Alzheimer's disease diagnosis from diffusion tensor images using convolutional neural networks. *PloS one, 15*, e0230409.
11. Sathiyamoorthi, V., Ilavarasi, A. K., Murugeswari, K., Ahmed, S. T., Devi, B. A., & Kalipindi, M. (2021). A deep convolutional neural network based computer aided diagnosis system for the prediction of Alzheimer's disease in MRI images. *Measurement, 171*, 108838.
12. Verma, V., Mishra, A. K., & Narang, R. (2019). Application of Bayesian analysis in medical diagnosis. *Journal of the Practice of Cardiovascular Sciences, 5*, 136–141.
13. Polotskaya, K., Muñoz-Valencia, C. S., Rabasa, A., Quesada-Rico, J. A., Orozco-Beltrán, D., & Barber, X. (2024). Bayesian Networks for the Diagnosis and Prognosis of Diseases: A Scoping Review. *Mach. Learn. Knowl. Extr., 6*, 1243-1262. URL: https://api.semanticscholar.org/CorpusID:270262599
14. Tan, W., Cao, Y., Ma, X., Ru, G., Li, J., Zhang, J., . . . Li, J. (2023). Bayesian Inference and Dynamic Neural Feedback Promote the Clinical Application of Intelligent Congenital Heart Disease Diagnosis. *Engineering, 23*, 90-102. doi:https://doi.org/10.1016/j.eng.2022.10.015
15. Polotskaya, K., Muñoz-Valencia, C. S., Rabasa, A., Quesada-Rico, J. A., Orozco-Beltrán, D., & Barber, X. (2024). Bayesian Networks for the Diagnosis and Prognosis of Diseases: A Scoping Review. Machine Learning and Knowledge Extraction, 6(2), 1243–1262. doi:10.3390/make6020058
16. Buathong, P., Wan, J., Astudillo, R., Daulton, S., Balandat, M. & Frazier, P.I.. (2024). Bayesian Optimization of Function Networks with Partial Evaluations, *Proceedings of the 41st International Conference on Machine Learning*, in *Proceedings of Machine Learning Research* 235:4752-4784 Available from https://proceedings.mlr.press/v235/buathong24a.html.
17. Krishnan, R., Subedar, M., & Tickoo, O. (2020). Specifying Weight Priors in Bayesian Deep Neural Networks with Empirical Bayes. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(04), 4477–4484. doi:10.1609/aaai.v34i04.5875
18. Kharoua, R. E. (2024). Alzheimer's Disease Dataset . *Alzheimer's Disease Dataset* . Kaggle. doi:10.34740/KAGGLE/DSV/8668279

Serhii Hladiholov **Сергій Гладіголов**	PhD student, Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, Ukraine, e-mail: hladiholov.s@gmail.com. https://scholar.google.com/citations?user=19aNHgkAAAAJ&hl=uk	аспірант кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, Україна.
Oleksii Kozachko **Олексій Козачко**	PhD, Associate Professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, Ukraine, e-mail: lekoz80@gmail.com https://scholar.google.com.ua/citations?user=VIrniUYAAAAJ&hl=uk	кандидат технічних наук, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, Україна.