ARTIFICAL INTELLIGENCE CONFERENCE

BOOK OF ABSTRACTS

BELGRADE, DECEMBER 26-27, 2024

PROGRAM COMMITTEE

Co-chairs:
Vladan Devedžić, SASA, Belgrade and Faculty of Organizational Sciences, University of Belgrade
Miodrag Mihaljević, SASA, Belgrade and Mathematical Institute, SASA, Belgrade
Dušan Teodorović, SASA. Belgrade

Members:
Sonja Dimitrijević, "Mihajlo Pupin" Institite, Belgrade
Mirko Dinulović, Faculty of Mechanical Engineering, University of Belgrade
Igor Đurović, Montenegro Academy of Sciences and Arts, Podgorica, Montenegro
Mirjana Ivanović, Faculty of Sciences, University of Novi Sad
Predrag Janičić, Faculty of Mathematics, Univeristy of Belgrade
Jelena Jovanović, Faculty of Organizational Sciences, University of Belgrade
Aleksandar Kostić, SASA, Belgrade
Danijela Milošević, Faculty of Technical Sciences in Čačak, University of Kragujevac
Marija Mitrović Dankulov, The Institute of Physics Belgrade
Boško Nikolić, School of Electrical Engineering, University of Belgrade
Zoran Ognjanović, Mathematical Institute SASA, Belgrade
Vladimir Srdanović, Mathematical Institute SASA, Belgrade
Ranka Stanković, Faculty of Mining and Geology, University of Belgrade
Leonid Stoimenov, Faculty of Electronic Engineering, Univeristy of Niš
Slobodan Vukosavić, Faculty of Electrical Engineering, University of Belgrade and SASA, Belgrade

ORGANIZING COMMITTEE

Co-chairs:
Tatjana Davidović, Mathematical Institute SASA, Belgrade
Andreja Tepavčević, Mathematical Institute SASA, Belgrade and Faculty of Sciences, University of Novi Sad

Members:
Ana Anokić, Academy of Technical and Art Applied Studies, Belgrade
Slobodan Jelić, Faculty of Civil Engineering, Univeristy of Belgrade
Dragan Jočić, Mathematical Institute SASA, Belgrade
Una Stanković, Mathematical Institute SASA, Belgrade
Vanja Stepanović, Faculty of Agriculture, University of Belgrade
Nataša Stević, SASA, Belgrade
Biljana Stojanović, Mathematical Institute SASA, Belgrade
Petar Vranić, Mathematical Institute SASA, Belgrade
Dejan Vukelić, Mathematical Institute SASA, Belgrade
Božo Vuković, Mathematical Institute SASA, Belgrade
Anđelka Zečević, Mathematical Institute SASA, Belgrade
Anastazia Žunić, Mathematical Institute SASA, Belgrade

ORGANIZERS
Department of Technical Sciences SASA, Belgrade
Mathematical Institute SASA, Belgrade
Academy Committee for Artificial Intelligence, SASA, Belgrade

Upgrading the Alzheimer's Disease Prediction Model Using Generative Artificial Intelligence

Halida Karišik[1], Aldina Avdić[1], Ulfeta Marovac[1]

[1]Department of Technical and Technological Sciences, State University in Novi Pazar, halidakarisik6@gmail.com, apljaskovic@np.ac.rs, umarovac@np.ac.rs

Abstract. Generative Artificial Intelligence (GAI) offers significant opportunities in machine learning by enabling the generation of synthetic data that reflect the key characteristics of original datasets. This approach is particularly valuable in domains where collecting sufficiently large and high-quality data is challenging, such as in predicting diseases like Alzheimer's. In this research, GAI is employed to generate synthetic brain images and tabular data from Alzheimer's patients, aiming to improve the accuracy and performance of machine learning models.

Data augmentation has emerged as a powerful strategy for enhancing the performance of machine learning algorithms [1, 2]. By applying transformations such as rotation, flipping, and cropping, new synthetic samples are generated, particularly in the field of computer vision [3]. In the medical domain, data augmentation plays a crucial role due to the limited availability of data [4, 5]. This approach not only increases the diversity of training datasets but also helps improve model robustness.

The aim of this research is to examine the role of generative artificial intelligence in improving models for predicting Alzheimer's disease by generating synthetic brain images and tabular data. The focus of the research is on analyzing how synthetic data can contribute to increasing the accuracy and robustness of machine learning models, utilizing methods such as logistic regression, the k-nearest neighbors algorithm, neural networks, and generative adversarial networks (GANs). The goal is also to assess the advantages and limitations of this approach in various aspects of prediction.

The research relies on two publicly available data sets: one tabular and one image set. The tabular data set for Alzheimer’s disease detection contains health information for 2,149 patients, including demographic details, lifestyle factors, medical history, and disease diagnosis. The MRI image set includes 5000 images categorized into two classes: Demented and Non-Demented. These images were manually collected and carefully labeled to ensure accuracy.

Synthetic tabular data successfully replicates the key patterns of the original dataset, though there are minor variations in extreme values. Different classification models demonstrated varying performance depending on the percentage of synthetic data used. Logistic regression achieved the best F1-score of 78% with 0% synthetic data, while K-nearest neighbors had the best result of 35% with 80% synthetic data. Neural networks showed a best F1-score of 71% with 20% synthetic data, and both naive Bayes and Random Forest performed best with the original data, achieving F1-scores of 76% and 90%, respectively.

On the MRI image dataset, a comparison between image augmentation techniques and GAN-generated images was performed, and it was found that the GAN-generated images resulted in a better-performing model. GAN models have shown notable effectiveness with image data, significantly improving model performance by expanding the original data set with synthetic images, enabling better generalization to unseen data (Table 1.). This highlights the value of synthetic images in complex tasks like image classification, where data diversity and quantity are essential for high performance.

In contrast to the success observed with image data, the application of Generative Adversarial Networks (GANs) to tabular data has not proven to be as effective. The performance of models using synthetic tabular data is less consistent, often showing limited improvement and varying significantly based on the proportion of synthetic data utilized. While GANs can still be useful in scenarios with insufficient data, their ability to enhance model performance is significantly lower than that achieved with image datasets, indicating that

the effectiveness of GAN-generated data is highly contingent upon the type of data and the specific methodologies applied.

Table 1: Performance Results of the Model Based on the Percentage of Synthetic Images

Percentage of Synthetic Images	Accuracy	Sensitivity	Precision	F1-score
0%	70.77%	70.77%	70.46%	70.08%
50%	76.65%	76.65%	77.93%	76.84%
100%	73.62%	73.62%	74.97%	73.83%

Additionally, the use of GAN-generated images can address challenges associated with limited datasets in specific domains, particularly in medical imaging. Research has shown that incorporating synthetic images not only improves model accuracy but also helps reduce overfitting, as models trained with varied data distributions are more likely to generalize effectively to real-world scenarios. Consequently, GANs have emerged as a powerful solution for tackling data scarcity while delivering robust results in machine learning applications.

Keywords: Generative adversarial networks, Disease prediction, Classification of images, Classification of tabular data

References:

[1] Oza, P., Sharma, P., Patel, S., & Kumar, P. (2023). Computer-aided breast cancer diagnosis: Comparative analysis of breast imaging modalities and mammogram repositories. *Current Medical Imaging*, *19*, 456–468.

[2] Ma, F., Wang, Y., Gao, J., Xiao, H., & Zhou, J. (2020). Rare disease prediction by generating quality-assured electronic health records. In *Proceedings of the 2020 SIAM International Conference on Data Mining* (pp. 514–522). Society for Industrial and Applied Mathematics.

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144.

[4] Kong, Q., Tong, B., Klinkigt, M., Watanabe, Y., Akira, N., & Murakami, T. (2019, July). Active generative adversarial network for image classification. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 4090–4097).

[5] Alauthman, M., Al-Qerem, A., Sowan, B., Alsarhan, A., Eshtay, M., Aldweesh, A., & Aslam, N. (2023, March). Enhancing small medical dataset classification performance using GAN. In *Informatics*, *10*(1), 28.

Table of Contents: