Using GANs and MLP Artificial Neural Networks to support early diagnosis of Alzheimer's disease: a study on the potential of artificial data expansion

Jonathan Bandeira
Escola Politécnica de Pernambuco (POLI)
Universidade de Pernambuco (UPE)
Recife, Brazil
jsb2@ecomp.poli.br

Mêuser Valença
Escola Politécnica de Pernambuco (POLI)
Universidade de Pernambuco (UPE)
Recife, Brazil
meuservalenca@gmail.com

Renan Costa Alencar
Escola Politécnica de Pernambuco (POLI)
Universidade de Pernambuco (UPE)
Recife, Brazil
rca2@ecomp.poli.br

***Abstract*— The life expectancy of the population in the most developed countries is growing every day and, consequently, there is an increase in various age-related diseases. In Brazil, just over 1.1 million people have Alzheimer's disease (AD). In 2019, according to the World Health Organization, Alzheimer's disease and other dementias were the third leading cause of mortality in the Americas and Europe. Despite being a degenerative and irreversible disease, if diagnosed early, treatments can be performed to slow the progression of symptoms and ensure a better quality of life for the patient. Most papers that study Computational Intelligence solutions to support diagnosis follow an approach based on neuroimaging evidence. In addition to this, another approach that has been gaining prominence is biomolecular analysis. Following this approach, Ray et al., Ravetti & Moscato and Dantas & Valença performed studies with classifiers based on the statistical area or on Computational Intelligence to support the early diagnosis of the disease. The work was carried out from a data set with values of 120 blood proteins. With this, they were able to classify whether the patient could be diagnosed with AD. This work aimed to use a traditional approach with a proposed Multilayer Perceptron Artificial Neural Network model to perform the early diagnosis of a patient with or without AD and compare the results obtained with the results of the related works mentioned. In addition, this work had, as its main objective, the evaluation of potential use of synthetic data generated through a Generative Adversarial Network (GAN) in the training and testing of the proposed classification model. The proposed MLP model presented mean rates of 94% accuracy, 100% sensitivity and 87% specificity. These results showed the remarkable capacity of the model, especially in recognizing the sick class. As for the synthetic data generated, these showed potential in optimizing the performance of the original model through artificial data expansion.**

Keywords—Alzheimer's disease, Multilayer Perceptron, Artificial Neural Networks, Generative Adversarial Networks, data augmentation.

I. INTRODUCTION

In the last decades, life expectancy has increased. Because of this, more challenges for the elderly part of the population arise [1]. One of the main concerns is the development of Chronic Non-Communicable Diseases (NCDs) in older adults. Among these stands out dementias such as Alzheimer's disease (AD). AD is a disease that already affects more than 47 million worldwide [2].

Dementia declines an individual's cognitive functions, which directly affect their behavior and quality of life. Dementias can be reversible or irreversible. Irreversible dementias can also be degenerative and progressive, which get worse over time [3]. In this field, we can highlight a range of conditions from a mild impairment, such as Mild Cognitive Impairment (MCI), to diagnoses of severe dementia conditions such as Alzheimer's Disease, discovered in 1907 by Alois Alzheimer's.

Alzheimer's Disease is a progressive neurodegenerative disorder that causes irreversible damage to an individual's brain, deteriorating their cognitive capacity and memory (the ability to recall old information or learn new information) and compromising their daily routine and behavior. AD is currently the form of degenerative dementia with the highest incidence globally, affecting mainly the elderly over 65 years of age in its late or senile manifestation. Despite this, it can also rarely affect young people at an early age.

According to the Brazilian Medical Association and the National Supplementary Health Agency, the worldwide prevalence of dementia becomes higher as age increases. In their last update, the following statistics were presented: 1.2% between 65 and 69 years old; 3.7% from 70 to 74 years old; 7.9% from 75 to 79 years old; 16.4% between 80 and 84 years old; 24.6% from 85 to 89 years old; 39.9% from 90 to 94 years old; 54.8% from 95 years onwards [2].

Considering the Brazilian elderly population of approximately 15 million people and the incidence of dementia in Brazil, it is estimated that the scope of dementia in the country reaches a value of approximately 1.1 million individuals [4].

As stated by the American Alzheimer's Association, the situation is even more alarming in other countries, such as the USA. In 2016, estimates were of 5.4 million individuals with AD, 11% of whom were aged 65 and over 32% of those aged 85 and over [5]. Projections made this year for the figures in 2050 stated that 51% of the population over 65 years old would be affected by the disease. Also, around 7 million people over the age of 85 will be affected [5].

Along with the World Health Organization, in the last two decades, between 2000 and 2019, Alzheimer's disease and other forms of dementia ranked the top ten diseases responsible for global mortality [6]. In this ranking, AD and other dementias appeared as the third leading cause of mortality in the Americas and Europe, in the survey concluded in 2019 [6].

It is important to emphasize that, like other chronic diseases, although irreversible, AD can be controlled if

diagnosed in its initial phase. This early diagnosis is crucial to delay degenerative progress and guarantee a better quality of life for the patient [2].

At present, the set of solutions with Computational Intelligence as support for diagnosis is based on three approaches: cognitive tests, neuroimaging evidence, and biomolecular tests.

The first approach is based on exams such as the Mini-Mental State Examination and has as its main positive point the most straightforward data to understand from questionnaires [2].

Neuroimaging is considered the current state-of-the-art of the three approaches and is based on evidence and imaging exams, such as CT scans, PET scans, and MRIs. The accuracy of the data that forms the basis used in the training of Machine Learning models is high. This high accuracy is due to the images of cerebral autopsies, in the post-mortem period of each patient, provide higher values of sensitivity and specificity. Papers that use this approach are the most common in the scientific community [7].

Definitively, biomolecular analysis is an approach that has gained notoriety in the last two years due to new evidence on the pathophysiology of the disease according to Dalmagro et al. [9]. The great discoveries in this area, such as a new high-precision and low-cost blood test capable of detecting the development of AD up to twenty years earlier [10]; and, more recently, advances in the studies of a vaccine have been capable of preventing the development of the disease [11]. This approach focuses on studying genetic, molecular, protein characteristics and other features present in blood plasma and are measured by laboratory tests [8].

Following the last approach mentioned, Ray et al. [8] developed a study using a dataset of 259 patients containing the concentration levels of 120 proteins contained in the blood plasma samples of these patients. Of these samples, 222 comprise samples of diagnoses of Alzheimer's Disease, Mild Cognitive Impairment, other dementias, and cases without dementia. This work concluded that a signature resulting from the combination of 18 proteins out of 120 allowed the early diagnosis of AD. Using the PAM technique for classification, an accuracy rate of 89% was shown. Also, 90% of correct answers were obtained in positive cases (recall), and 88% of correct answers in negative cases (specificity) in the test set of AD.

Using the same dataset cataloged by Ray et al., Ravetti & Moscato [12] used classification algorithms available in the Weka software and obtained an average accuracy rate of 93%. 96% of correctness were achieved in positive cases (recall), and 90% of correctness in negative cases (specificity) for the tests set of AD. This paper considered a better combination of 5 proteins.

In 2013, Dantas & Valença [13], another study that used the same database, used a Reservoir Computing (RC) framework, and obtained, on average, 94.34% accuracy for the tests set of AD. It considered the same combination of 5 proteins that were used by Ravetti & Moscato.

To select a set of proteins, Ray et al. [8] and Ravetti & Moscato [12] used the already cited PAM tool to extract characteristics. The selection made by Ravetti & Moscato resulted in a subset of the selection already performed by Ray et al. [12]. On the other hand, Dantas & Valença [13], used the Random Forests technique with information gain analysis (based on entropy) to find the selection, which coincided with the selection of 5 proteins by Ravetti & Moscato. This work used this same selection of 5 proteins to carry out the experiments.

The present work used an Artificial Neural Network (RNA) Multilayer Perceptron (MLP) model in the same 5-protein signature used by the two most recent works cited. The performance was compared with works in the literature and classifiers in their default settings, coming from the SciKit-Learn library written in Python. After that, a Generative Adversarial Network model was used to generate 10,000 synthetic samples for testing in two scenarios: training on real samples with testing on synthetic samples, and training on synthetic samples with testing on real samples.

II. Machine Learning techniques

A. *SciKit-Learn library algorithms*

Python language was used to implement and import necessary machine learning algorithms. The following algorithms were used from the SciKit-Learn library [14]: Random Forest, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines. For all of these, the configuration used was the standard provided by the library. The proposed classification model was a Multi-Layer Perceptron Neural Network, based on the sequential model provided by Python's Keras API [15].

B. *Artificial Neural Networks*

Artificial Neural Networks (ANN) are computational models inspired by the biological central nervous system, particularly brain neurons, capable of machine learning and pattern recognition. Another way to define this technology is to state that an ANN is a function approximator. The primary purpose of an ANN is to simulate the connectionist behavior of a chain of neurons so it can learn from the environment and improve its performance [16].

Computationally, the objective of an ANN is to reduce forecasting errors by minimizing a given function, called a cost function. The term cost function is used to determine the forecasting errors of the output neurons over a cycle or time. The calculation performed for a single example is called a loss function [17].

A training procedure is carried out for the loss function. An iterative adjustment process is applied to the weights of connections (synapses) between neurons within an ANN [16]. When a neural network can generalize a solution to a specific class of problems, it is said that there was learning [16].

The learning process of an ANN is verified by changing the weights. The initial weights, usually generated at random, are modified iteratively by a training algorithm that follows one of the following paradigms [17]:

- Supervised Learning: a training set is presented, consisting of the inputs and desired corresponding outputs.
- Reinforcement Learning: for each input presented, an indication (reinforcement) about the adequacy of the corresponding outputs produced by the network is produced.
- Unsupervised Learning: the network has its weights adjusted without the use of desired input-output pairs nor indications about the adequacy of its corresponding outputs produced by the network.

To design and set an ANN, we must know some characteristics like the neuron features, the network topology, and the training rules. A well-designed and configured ANN brings in addition to learning some other benefits, such as high power of generalization, flexibility, fault tolerance, self-organization, and parallel information processing [17].

In practice, ANN can be applied to solve problems of function approximation, classification, pattern recognition, and time-series prediction [17].

Multilayer Perceptron is part of an ANN's class architectures known as feedforward, whose neurons are grouped in layers. The signals travel through the network in a single direction, from input to output, and neurons in the same layer are not interconnected [17]. The backpropagation algorithm is usually used to train an MLP model [16]. Fig. 1 shows the architecture of an MLP neural network.

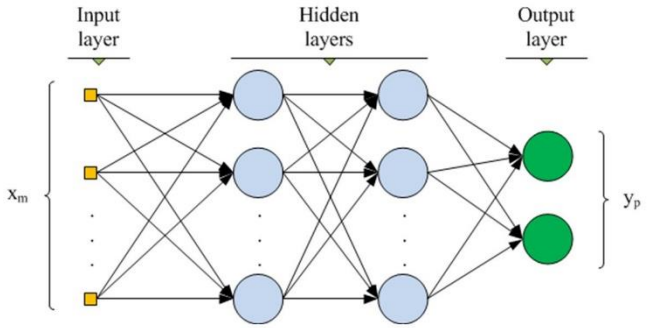

Fig. 1. Multilayer Perceptron Neural Network architecture

C. *Generative Adversarial Networks*

The second proposed technique was GAN, a framework that represents a new class of machine learning models. This framework was developed by Goodfellow et al. [18], which consists of two Artificial Neural Networks competing against each other in a sort of game.

In this game, one of the neural networks assumes the role of the generator while the other has the function of discriminator [18]. The generator G is based on a given training set of real data, and with the addition of noise, it learns to produce new data like the real ones and with the same statistics as the original ones [18]. The new data generated is called synthetic data. The discriminator D, in turn, tries to distinguish real from synthetic instances. Fig. 2 shows the organizational structure of a GAN.

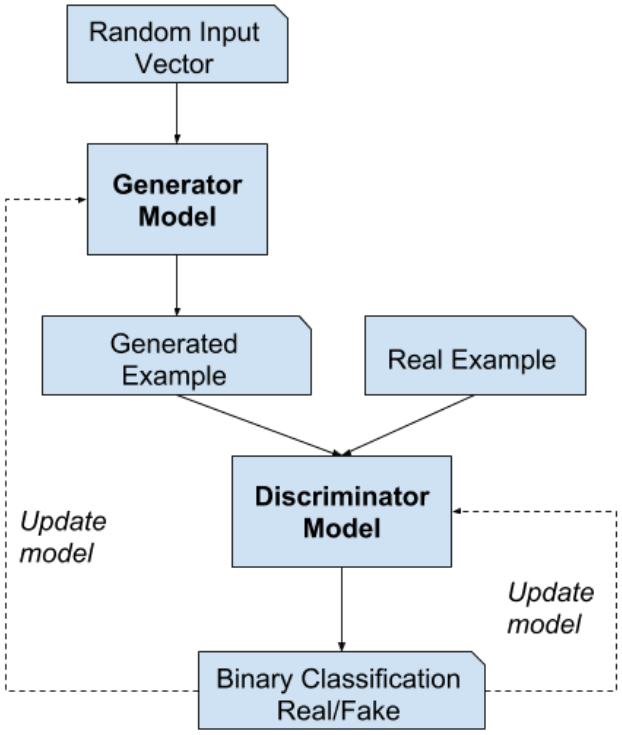

Fig. 2. GAN model architecture

Competitive learning is possible through the loss rule defined for this process and the way neural networks are trained in the framework. The loss rule involves maximizing and minimizing two logarithmic functions, given by (1) [18].

$$Min_G Max_D VGAN\ (D,G) = E_{x\sim Pdata(x)}[\log(D(x))] + E_{z\sim Pz(z)}[\log(1-D(G(z)))] \quad (1)$$

In $Min_G Max_D VGAN\ (D,G)$, Min_G is the generator loss minimization given only by the second term of the equation, while Max_D is the discriminator loss maximization given by all the equation. The term $VGAN\ (D,G)$ is where the maximum and minimum functions apply, which can be described as the loss vector of the GAN model, where D corresponds to the discriminating model and G is the generating model.

The discriminator D tries to maximize both functions and maintain good performance to identify the real data as true and identify the synthetic data as false [17]. On the other hand, the generator G tries to minimize the second function, generating data increasingly like the real ones, to reduce the performance of the discriminator in both situations (real and synthetic data) [18]. That way, the discriminator will not be able to specify what is real or fake.

In the first log loss function (log), given by $E_{x\sim Pdata(x)}[\log(D(x))]$, we can highlight the discriminator applied to real data and distributions (x~Pdata(x)) in the form D(x), where E is the loss value calculated in $[\log(D(x))]$.

In the calculation of the loss by the second equation, given by $E_{z\sim Pz(z)}[\log(1 - D(G(z)))]$, the discriminator is applied to the synthetic data and distributions generated in the form D(G(z)), where G(z) is the data generated by the generator G and (z~Pz(z)) is the real data and distributions imbued with noise.

Both functions of the equation are complementary, so the second loss function (E) is calculated by the function $[\log(1 - D(G(z)))]$. Thus, while the first function is D(a), the second is (1 – D(a)). In the first function, this argument "a" is the real data x, and in the second, the generated data G(z).

III. Materials and Methods

A. Dataset

The database used in the development of this work was the same developed by Ray et al. [8] in their publication. It contains the values of 120 proteins found by analyzing blood samples from 222 patients diagnosed with AD, MCI, other dementias, or no dementia. The base's goal is to classify whether a patient can be diagnosed positively or negatively concerning Alzheimer's Disease (AD or NAD, respectively).

Initially, Ray et al. [8] subdivided the dataset into three sets. The first one contained 83 blood samples with values of 120 proteins. The second set contained information from 92 patients diagnosed with or without AD. On the other hand, the third contained 47 samples from patients diagnosed with MCI that progressed over time to AD or other dementias or maintained the final diagnosis of MCI after a few years.

In this analysis, only the training sets and AD tests were considered. Diagnoses related to other dementias were also excluded, having been worked on instead of 92 instances of tests, only 81 instances. It contains only the cases of patients with AD or without dementias. Table I shows which of the 120 proteins were contained in the signatures defined by Ravetti & Moscato [12] and Dantas & Valença [13], also used in this work.

TABLE I. Description of the protein signature used

Number of proteins in signature	Proteins
5	**IL-1α, IL-3, EGF, TNF-α and G-CSF**

Three datasets were used in this work:

- A dataset with 83 samples used for training the classifiers in the first and second experiments;
- A dataset with 81 samples used for tests, i.e., performance evaluation of the classifiers in the first and second experiments;
- A synthetic dataset generated by GAN with 10,000 samples used for both training and testing in the second experiment.

The input data have been normalized, i.e., transformed to be contained in a single numerical range in these datasets. This normalization is essential, considering that the different variations of each feature could skew the training and cause a loss in the ability to generalize the models.

The input data were normalized using the MinMaxScaler function of the SciKit-Learn library. The range used as the default, between 0 and 1. The cited function uses the formula given by (2):

$$X_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Equation $X_{normalized}$ corresponds to the new value of a given instance of a characteristic after normalization. The x_{max} value is the highest value among all instances of a given characteristic. The value x_{min} is the smallest value among all instances of a given characteristic. Finally, the value of x corresponds to the original value of a given instance of a characteristic before it was normalized.

The labels were transformed into numerical information, assuming the values 0 (zero) for positive class and 1 (one) for negative class.

In the original dataset [8], a significant number of outliers were identified within all sets. Because of the division carried out by Ray et al., outliers were equally distributed in both sets of training and testing, and this gives the idea that the instances are equally representative. However, using a smaller fraction of data for testing or when using a more significant number of instances in training, it is possible to notice that these outliers are primarily found in instances of positive cases. These outliers can be seen in the synthetic data generated. In practice, this means that instances of negative cases are more representative than the other ones.

B. Experiments

We chose to use an MLP as it is the simplest ANN model capable of solving this classification problem. As the problem is not linearly separable, elementary models are not viable.

More robust models using deeper architectures (Deep Learning) were not tested due to their high computational cost. It is necessary to verify whether the gain in this operation is worth, that is, if the cost x benefit of these models would be better than those of using an MLP.

Regarding the use of recurrent ANN models, this study has already been carried out in Dantas & Valença [13]. It has already been proved that the cost x benefit of using a recurrent model is lower than using an MLP.

Thus, in this study, we chose to use an MLP model with a good proved cost-benefit to solve this problem and to seek performance improvements through tests with artificial data expansion using GAN models for this purpose.

As already described, the classification models from the SciKit-Learn library had their default settings maintained. The MLP model, based on a sequential Keras API model, presented configurations as shown in Table II.

TABLE II. MLP MODEL CONFIGURATION

Parameters	Value of parameters for MLP model
Number of neurons in the input layer	5
Number of hidden layers	2
Number of neurons in the first hidden layer	120
Number of neurons in the second hidden layer	60
Number of neurons in the output layer	1
Activation function on hidden layers	ReLU
Activation function in the output layer	Sigmoid
Optimizer	Adam
Optimizer parameters	Learning rate = 0.0002; beta 1 = 0.5; beta 2 = 0.999; epsilon = 0.0000001.
Cost function	Binary cross entropy
Epochs	100
Batch size	10

The GAN model was also implemented using the Keras library and had its configuration performed as shown in Tables III and IV. Table III refers to the Generator model, and Table IV is referring to the Discriminator model. The GAN model was trained using all 164 considered AD (83 training instances and 81 test instances).

TABLE III. GAN GENERATOR CONFIGURATION

Parameters	Parameter values for layers in common for both branches	Parameter values for the categorical generation branch	Parameter values for the numerical generation branch
Number of hidden layers	2	2	2
Number of neurons in the first hidden layer	8	32	64
Number of neurons in the second hidden layer	16	64	120
Activation function on hidden layers	LeakyReLU	LeakyReLU	LeakyReLU
Number of neurons in the output layer	-	2	120
Activation function in the output layer	-	Softmax	Softmax
Latent dimension	80	-	-
Other layers	BatchNormalization (momentum = 0.8)	BatchNormalization (momentum = 0.8)	BatchNormalization (momentum = 0.8)

TABLE IV. GAN DISCRIMINATOR CONFIGURATION

Parameters	Value of parameters for Discriminator model
Number of neurons in the input layer	122 (alternating between actual values and generated values)
Number of hidden layers	5
Number of neurons in the first hidden layer	128
Number of neurons in the second hidden layer	64
Number of neurons in the third hidden layer	32
Number of neurons in the fourth hidden layer	16
Number of neurons in the fifth hidden layer	8
Number of neurons in the output layer	1
Activation function on hidden layers	LeakyReLU
Activation function in the output layer	Sigmoid
Optimizer	Adam
Optimizer parameters	Learning rate = 0.0002; beta 1 = 0.5; beta 2 = 0.999; epsilon = 0.0000001.
Cost function	Binary cross entropy

The first experiment involved training and testing all classifiers with only the original database sets with the selection of 5 proteins. The second experiment, on the other hand, also used the generated synthetic database. In the end, the performances of the other classifiers were compared with the model proposed in both simulations. Finally, the result of this work is compared to the other related works cited.

For each algorithm, a seed was defined to guarantee the reproducibility of the results. For some algorithms from the SciKit-Learn library (Random Forest and SVM), the random_state variables of these algorithms were used. This variable was defined as 0 (zero). For the Keras MLP model, NumPy.random.seed and TensorFlow.random.set_the seed were imported. Both were assigned the value 7 (seven).

Only the performances of the test set were considered to compare models. The comparison considered the accuracy metrics to ascertain the overall performance of each model: first the sensitivity, to measure the performance of each model in terms of predicting the positive class; and then, the specificity, aiming to measure the performance of each model regarding the negative class prediction.

The percentages shown in the comparative results tables of the models' performances are the average values resulting from thirty runs in different resampling of the test set.

When comparing this work with related works, the comparison methodology is similar, and the percentages also correspond to the average of the results in the test set.

IV. RESULTS

After the execution of the first experiments, the results presented in Table V were obtained.

TABLE V. COMPARISON BETWEEN RESULTS OF THE PROPOSED MODEL AND THE OTHER CLASSIFIERS

Techniques	Accuracy	Recall	Specificity
K-NN	**90%**	**100%**	**79%**
Naïve Bayes	**81%**	**98%**	**64%**
SVM	**88%**	**100%**	**74%**
Random Forest	**81%**	**98%**	**64%**
MLP (proposed model)	**94%**	**100%**	**87%**

The MLP Neural network presented the most balanced results, considering all metrics, in the first experiment. Finally, Table VI shows that the proposed model results were similar to the results of all related works cited, emphasizing the sensitivity rate of the model.

TABLE VI. COMPARISON BETWEEN THE PERFORMANCES OF THE MODEL OF THE PRESENT WORK WITH THE RELATED WORKS

Works	Accuracy	Recall	Specificity
Ray et al. (2007)	**89%**	**90%**	**88%**
Ravetti & Moscato (2008)	**93%**	**96%**	**90%**
Dantas & Valença (2013)	**94.34%**	-	-
This work	**94%**	**100%**	**87%**

For the second experiment, Tables VII and VIII present the results obtained considering the training performed on synthetic data with tests on synthetic data and tests on real data. Table IX shows the results from training with the real data set and the test on a synthetic dataset.

TABLE VII. PERFORMANCE RESULTS OF TRAINED AND TESTED CLASSIFIERS USING SYNTHETIC DATASET

Techniques	Accuracy	Recall	Specificity
K-NN	**87%**	**92%**	**80%**
Naïve Bayes	**80%**	**86%**	**72%**
SVM	**88%**	**93%**	**81%**
Random Forest	**88%**	**91%**	**82%**
MLP (proposed model)	**87%**	**91%**	**82%**

TABLE VIII. PERFORMANCE RESULTS OF CLASSIFIERS TRAINED ON SYNTHETIC DATASET AND TESTED ON REAL DATASET

Techniques	Accuracy	Recall	Specificity
K-NN	**86%**	**86%**	**87%**
Naïve Bayes	**83%**	**69%**	**97%**
SVM	**83%**	**71%**	**95%**
Random Forest	**83%**	**81%**	**85%**
MLP (proposed model)	**84%**	**74%**	**95%**

TABLE IX. PERFORMANCE RESULTS OF CLASSIFIERS TRAINED ON REAL DATASET AND TESTED ON SYNTHETIC DATASET

Techniques	Accuracy	Recall	Specificity
K-NN	**75%**	**95%**	**48%**
Naïve Bayes	**75%**	**94%**	**48%**
SVM	**76%**	**95%**	**49%**
Random Forest	**72%**	**94%**	**40%**
MLP (proposed model)	**78%**	**90%**	**62%**

In the first case, it is possible to observe that the synthetic data has a similar distribution to the real data since the training and tests in real data present a behavior similar to the training and tests in synthetic data. Here, two subsets of 5,000 samples each out of a total of 10,000 samples were used.

The second case comprises training on real data, and tests on synthetic data impact the recognition of negative cases. Thanks to how the data was partitioned in the real training set, the outliers present in the real set make the model classify instances of healthy patients as positive AD cases.

In the third and last case of this experiment, the training was carried out with a subset of 5000 synthetic data and then test the real data set for tests. In this case, the opposite of the second case occurred. Here, the hit rate for negative cases was higher than the hit rate for positive cases. This variation was because samples from healthy patients are more representative than the other ones. Since the model was trained in a more significant number of samples, with greater access to samples in this class, this made it recognize more instances of this class to the detriment of the positive class instances.

V. CONCLUSIONS

Alzheimer's disease is currently one of the three most significant causes of mortality globally due to its complications. This pathology is becoming more and more present in society as the population's longevity increases. The disease is a progressive neurodegenerative disorder that is irreversible. Although there is still no cure, existing treatments can provide higher quality and life expectancy for patients diagnosed early, whether pharmacological or behavioral.

In this way, this work aimed to verify the impacts and future potentials arising from synthetic data in improving the performance of the more traditional classifiers for this problem. We also sought to investigate the potential of the proposed model concerning other classifiers and related works.

The proposed model showed results like those presented in the related works, emphasizing its ability to predict positive cases.

Furthermore, synthetic data has shown excellent potential in optimizing the model's performance with the combination of real data and synthetic data. This potential would enable greater capacity to generalize the model and improve the training of an ANN model, which generally requires a significant amount of data.

In the more practical scope of applications, we can highlight the potential to prove that GAN models can be useful even in the expansion of highly sensitive data (such as health data), i.e., it has wide applicability during data expansion for the development of critical systems. Moreover, it is crucial to mention the possibility of developing a specialist system to aid in the early medical diagnosis of neurodegenerative diseases such as Alzheimer's disease. This specialist system would be based on models presented in this work and the approach designed by Ray et al. [8].

To guarantee the reliability of the similarity of the generated data, it would be interesting, as a proposal for future work, to generate the data with a Conditional Generative Adversarial Network (CGAN) model [19], as these are based on a given condition to perform the sample generation. The condition used could be the sample labels, assuring that the data generated would be more faithful to the proposed labels.

Another proposal for relevant future work would be to treat the outliers of the dataset, which could certainly add to improvements in the performance of the models.

REFERENCES

[1] V.M.H. Lima, "Fatores Sociodemográficos do envelhecimento e as universidades da terceira idade", in VI Congresso Nacional em Educação (CONEDU), 2018.

[2] D. Apolinário, L.M.Q. Araújo, M.L.F. Chaves, L.C. Lopes, I.H. Okamoto, A.M. Ramos, A.T. Stein, N.C. Andrada, "Doença de Alzheimer: Diagnóstico", in Diretrizes Clínicas na Saúde Suplementar: Associação Médica Brasileira e Agência Nacional de Saíde Suplementar. Janeiro, 2011. Available in: < https://diretrizes.amb.org.br/ans/doenca_de_alzheimer-diagnostico.pdf> Last access: 2021-05-27.

[3] J. Gallucci, M.G. Tamelini, O.V. Forlenza, "Diagnóstico diferencial das demências", in Revista Psiquiatrica. Clínica. Vol.32, Número 3, São Paulo, Maio/Junho, 2005.

[4] Comissão Nacional de Incorporação de Tecnologias no SUS (CONITEC), "Protocolo clínico e diretrizes terapêuticas para Doença de Alzheimer", Relatório de Recomendação nº 285. Julho, 2017. Available in: <http://conitec.gov.br/images/Relatorios/2017/Recomendacao/Relatorio_PCDTDoen%C3%A7a_de_Alzheimer_267_17_final_SEC1207.pdf >. Last access: 2020-12-12.

[5] Alzheimer's Association, "Alzheimer's Disease Facts and Figures," in Alzheimer's Dement. 2016;12(4):459-509.

[6] World Health Organization(WHO), "WHO reveals leading causes of death and disability worldwide: 2000-2019," World Health DA Organization, Switzerland, 2020-12-09. Available in: <https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019>. Last access: 2020-12-20.

[7] I.R.R. Silva, G.S.L. Silva, R.G. de Souza, W.P. dos Santos, R.A.A. Fagundes, "Model based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease," in Neural Networks International Joint Conference on Neural Networks (IJCNN). Budapest, Hungary. IEEE, 14-19 July, 2019.

[8] S. Ray, M. Britschgi,C. Herbert,Y. Takeda-Uchimura, et al."Classification and Prediction of Clinical Alzheimer's Diagnosis based on plasma signaling proteins", NatMed 13: 1359–1362, 2007.

[9] A.P. Dalmagro, C.A. Cazarin, F.S. Zenaide, C.A. Dalmagro, "An update about the biochemical and molecular basis of Alzheimer's Disease" , in Brazilian Applied Science Review, Curitiba, V. 4, N. 1, P.118-130 Jan/Fev, 2020.

[10] S. Palmqvist, S. Janelidze, Y.T. Quiroz, et al. "Discriminative accuracy of Plasma Phospo-Tau217 for Alzheimer disease VS other Neurodegenerative disorders", in JAMA. 2020. 324(8):772–781.

[11] H., Davtyan A. Hovakimyan, S. Kiani Shabestari, et al. "Testing a MultiTEP-based combination vaccine to reduce Aβ and Tau Pathology in Tau22/5xfad", in Bigenic Mice. Alz Res Therapy, 11, 107 (2019).

[12] MG Ravetti, P. Moscato, "Identification of a 5-protein biomarker molecular signature for predicting Alzheimer's Disease", in PLOS ONE 3: E3111, 2008.

[13] L. Dantas, M. Valença, "Utilizando *Reservoir Computing* para realizar diagnóstico precoce de Mal de Alzheimer", in: BRICS SBIC, 2013.

[14] O. Grisel, G. Lemaitre, J. du Boisberranger and Chiara Marmo, "The Scikit-learn Consortium", at Inria Foundation. Available in: < https://scikit-learn.fondation-inria.fr/home/>. Last access: 2021-05-27.

[15] F. Chollet, "The Keras API", Available in: <https://keras.io/api/>. Last access: 2021-05-27.

[16] M. Valença, "Fundamentos das Redes Neurais: exemplos em Java", 2ª Ed. Recife: Livro Rápido, 2009.

[17] S. Haykin, "Redes Neurais: Princípios e Práticas", 2ª Ed. São Paulo: Bookman, 2001.

[18] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, "Generative Adversarial Nets", in Proceedings of The 27th International Conference on Neural Information Processing Systems, Volume 2, PP. 2672–2680, 2014.

[19] M. Mirza, S. Osindero, "Conditional Generative Adversarial Nets," in NIPS Deep Learning Workshop, 2014.