Automated Classification of Alzheimer's Disease using Deep Neural Network (DNN) by Random Forest Feature Elimination

Conference Paper · August 2019

DOI: 10.1109/ICCSE.2019.8845325

CITATIONS
34

READS
171

3 authors, including:

Gökçen Çetinel
Sakarya University
36 PUBLICATIONS **146** CITATIONS

SEE PROFILE

Ali Manzak
Florida State University
21 PUBLICATIONS **369** CITATIONS

SEE PROFILE

Automated Classification of Alzheimer's Disease using Deep Neural Network (DNN) by Random Forest Feature Elimination

Dilek Manzak
Dept. of Electronics Engineering
Sakarya University
Sakarya, Turkey
dilekmanzak@gmail.com

Gökçen Çetinel
Dept. of Electronics Engineering
Sakarya University
Sakarya, Turkey
gcetinel@sakarya.edu.tr

Ali Manzak
Dept. of Electrical & Computer Engineering
University of Detroit Mercy
Detroit, U.S.A.
manzakal@udmercy.edu

***Abstract*—Determining Alzheimer's Disease (AD) in its early stages is very important to prepare proper care for the patient. In this study, we aimed to create fast and accurate automated classification system to determine AD with the minimum data collected from the patient. Magnetic Resonance Imaging (MRI) is widely used to diagnose AD. When the cost of the technique and risks of the procedures are considered, there is a need for different solutions. With the availability of neural network chips, it is even possible to build portable devices for Alzheimer's detection. We propose fast and successful method to detect Alzheimer using Deep Neural Network (DNN). To reduce the complexity of the algorithm, Random Forest method was used to eliminate some of the features. Success of Random Forest to eliminate the features and success of DNN to detect AD are discussed.**

Keywords— Alzheimer's Disease, Deep Neural Network, Feature Selection, Random Forest

I. INTRODUCTION

Alzheimer's Disease is the most common type of dementia usually seen in people over the age of 60. The risk of developing AD significantly increases with the age and after 65 the risk doubles in every 5 years. AD is a progressive and non-reversible brain disorder and the causes of the disease are still a mystery. Currently, there is no cure for AD and the procedures are mostly for reducing the improvement of the disease. With baby boomers reaching the age of the risk of developing AD, the economic and social effects are huge on society. For that reason, it is very important to determine the disease as early as possible.

AD is the 6th leading cause of death in the United States, and 5th leading cause of death among those age 65 and older. Between 2000 and 2015 number of deaths related to heart disease, stroke and prostate cancer has decreased however deaths from AD increased 123% [1]. This number alone gives the importance of the need the research in AD. Alzheimer's Disease Neuroimaging Initiative (ADNI) [2] is launched to fulfill the requirement of the need in research enabling the sharing of the data among researchers around the world.

Some of the early signs of the disease are memory loss, poor decision-making, mood and personality changes and this stage is called Mild Cognitive Impairment (MCI). Mild Cognitive Impairment is the transitional stage from healthy to Alzheimer's Disease. Mild Cognitive Impairment causes decline in cognitive abilities such as memory and thinking skills. The decline in those skills is small but noticeable. MCI patients have higher risk of developing AD, however only 15% of MCI patients develop AD [3]. Available tests for diagnosis of Alzheimer's Disease have limitations [4]. Certain AD diagnosis require severe cognitive deficit and autopsy confirmation. However MCI diagnosis is not certain and biomarkers development for early diagnosis of AD is critical [5].

MRI is an effective tool to diagnose AD. However, use of MRI is costly and not always feasible. Better, cheaper and practical solutions are needed to detect AD. In this paper, we used DNN method to detect Alzheimer in early stages. DNN can be implemented in a chip successfully using neural network chips currently available in the market. That will make it possible to implement detection devices for AD in near future. In this study Random Forest method was used to eliminate features and reduce the complexity of the system.

II. PROCEDURE

A. Data Collection

ADNI Database: ADNI Alzheimer's Disease Neuroimaging Initiative [2], is a longitudinal multicenter study designed to develop clinical, imaging, genetic, and biochemical biomarkers for the early detection and tracking of Alzheimer's disease (AD). The study began in 2004 with 400 Mild Cognitive Impairment (MCI) patients, 200 Alzheimer Disease (AD)

patients and 200 healthy elderly patients. This study's primary goal was to develop biomarkers as outcome measures for clinical trials and was called ADNI-1. In 2009 the study expanded with 200 elderly MCI patients and was called ADNI-GO. ADNI-2 and ADNI-3, and all phases started different times. The aim of the study was to examine biomarkers in earlier stages of disease. When ADNI-Go ended in 2011, ADNI-2 began with existing ADNI-1 and ADNI-GO and 150 elderly controls, 100 early MCI, 150 late MCI and 150 AD.

B. Data Cleaning

ADNI data includes 12749 data from 1737 patients. Although some of the patients have follow ups for 10 years, some only have 2 years of follow-ups. The measurements were repeated in every 6 months however some of the patients' follow up data were missing. The study aims to find the similarities during transitions of MCI to AD. During data cleaning, we deleted healthy patients as well as patients without diagnosis. Distribution of patients' diagnoses is given in the Table 1 and mean values of some of the features are given in Table 2.

Total 602 data taken from 202 patients are used in the new set. First, we tried to examine the statistical analysis of these patients. In the data set, which contains 100 different features, some features have too many missing values. The features with missing values of 50% and above were deleted from the data set in the first stage. 28 of these patients were diagnosed with dementia within 3 years after MCI diagnosis, and 139 were not diagnosed with dementia in any of the sequential measurements. Three of the remaining patients were diagnosed as normal after the diagnosis of MCI, and others remained at the MCI to Dementia border. A similar approach used for data cleaning in [6]. 28 patients were positive for developing dementia from MCI, and 139 patients were named as stable (dementia negative) for staying MCI.

TABLE I. DISTRIBUTING OF PATIENTS IN ADNI DATA

Diagnose (DX)	Frequency	Percentage	Total Percentage
Dementia	1732	13.6	13.6
Dementia to MCI	12	.1	13.7
MCI	3935	30.9	44.5
MCI to Dementia	372	2.9	47.5
MCI to NL	77	.6	48.1
NA	3840	30.1	78.2
NL	2669	20.9	99.1
NL to Dementia	3	.0	99.1
NL to MCI	109	.9	100.0
Total	12749	100.0	

TABLE II. MEAN VALUES OF SOME FEATURES OF THE ADNI DATA

Features	MCI Positive (Mean)	MCI Negative (Mean)
AGE	73.6±7.35	74.8±7.46
EDUCATION	15.47 ±2.86	15.48±3.28
APOE4	0.730	0.709
ADAS11	17.97	11.28
ADAS13	27.9	18.3
Hippocampus	5444.61±890	6485.64±1053
MMSE	24.15±3.69	26.81±2.51
RAVLT.immediate	23.35±7.2	31.2±10.9
RAVLT.learning.bl	3.12±2.6	3.83±2.47
RAVLT.forgetting.bl	5.31±1.9	4.81±2.5
RAVLT.perc.forgetting.bl	83.81±19.8	64.80±32.5
FAQ.bl	4.26±3.8	3.33±3.7
Ventricles.bl	50498.41 ± 18385.2	46087.4±25689.8
WholeBrain.bl	952507.7 ± 107188.8	985323.1±106156

III. FEATURE SELECTION

A. Decision Tree

Decision tree learning is one of the fundamental methods in machine learning [7]. In decision tree learning, the set on which the training is performed is divided into sub-clusters according to various characteristics. This process is called recursive partitioning and continuous until the repetition has no effect on the estimation. In general, the way data is delivered during data mining is as follows:

$$(x, Y) = ((x_1, x_2, \ldots x_n), Y) \quad (1)$$

According to (1) x_1 to x_n values are input of the system, and Y is the output that need to be obtained.

The advantage of decision tree learning algorithms is simple to understand and interpret. Data is available with very little processing compared to most alternative techniques. The preparation stage is shorter and simpler than the other alternatives. It can be used to process both numeric and categorical data. Most machine learning algorithms are useful for either numerical applications or classification problems. Decision tree learning can be used in both areas. It has low computational complexity: Because it is simple and fast, it can process large amounts of data in a short time and it is preferable when the amount of data increases compared to the alternative methods.

B. Random Forest Feature Selection

After the deletion of the data with 50% and above missing values in the features, Random Forest is applied, and the feature was taken into consideration by using the Gini feature selection. Selected features importance order was given in Figure 1.

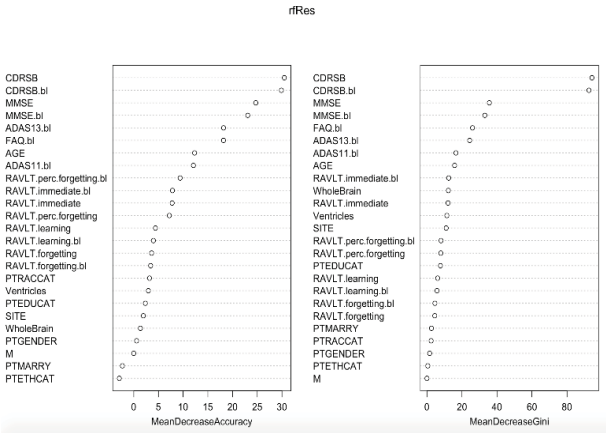

Fig. 1. Random Forest feature importance order

In machine learning applications, choosing features with high effect to output is very important to have an accurate model**.** Particularly in high-dimensional data, Random Forest is a method used in feature selection. Random Forest provides information on the severity of the properties, so it is possible to sort the properties according to high estimation rates.

C. Deep Neural Networks

Neural Networks are modelled after human brain and used to recognize patterns. Computations of neurons are sum of weighted synapses. Figure 2 shows diagram of neural networks [8]. The circles represent neurons and lines represent synapses, middle layer shows hidden layer, and output layer represents the class scores. Synapses take the input and multiply it by a weight. Neurons add the outputs from all synapses and apply an activation function. Neural Networks that have more than three layers and more than one hidden layer are called Deep Neural Networks [9].

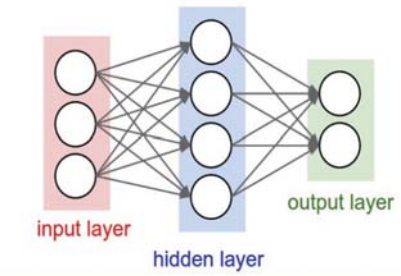

Fig. 2. Neurons and synapses

IV. RESULTS

Figure 3. shows Random Forest error change according to number of trees. Error rate drops when the number of trees increase, however after certain number the drop isn't well recognized.

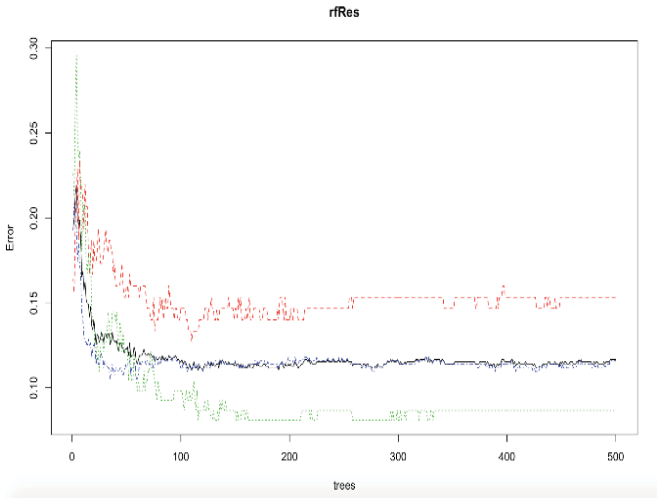

Fig. 3. Random Forest error change according to number of trees

Baseline measurements were chosen as training data, and test data was selected from 2nd year measurements. Following the application of Random Forest with 500 trees, the following features have been found to be of the utmost importance using the Gini significance feature; CDRSB, CDRSB.bl, MMSE, MMSE.bl, ADAS13.bl, FAQ.bl, ADAS11.bl, AGE, RAVLT.immediate.bl, RAVLT, Clinical Dementia Rating Scale (CDR), FunctionalActivities Questionnaires (FAQ) (AlzheimerASs Disease Assessment Scale), Mini-Mental State Examination (MMSE).

Using the features that were selected by Random Forest we created Neural Networks with second year measurements of the patients. Figure 3 shows the neural network error of the test data.

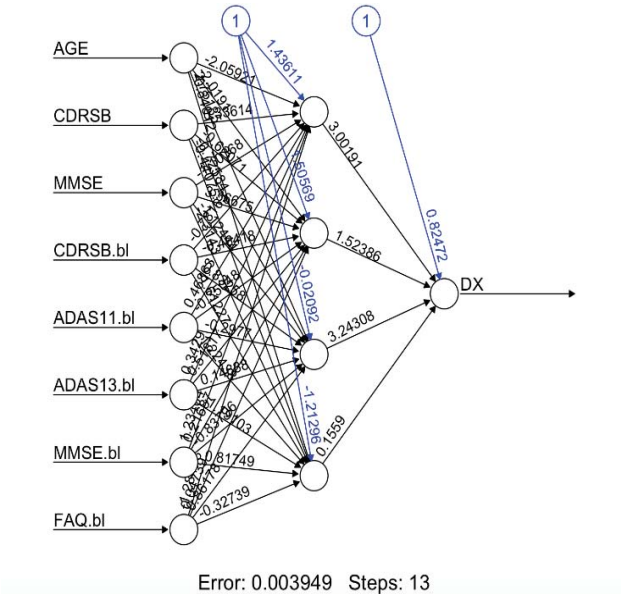

Fig. 4. Neural Networks

Table 3. shows confusion matrix of neural networks. 11 AD patients and 185 Normal Cognitive patients diagnosed correctly. The accuracy of the neural network is 67%.

TABLE III. CONFUSION MATRIX FOR NEURAL NETWORK

	AD	***CN***
AD	11	86
CN	9	185

After deciding which features are the most relevant we used selected features to train Deep Neural Network (Fig. 4). A Deep Neural Network is a Neural Network with more than two hidden layers of neurons. The number of layers were chosen as in [10]. According to deep neural network prediction, the confusion matrix is created in Table 3. The accuracy is calculated as 67%. When we used all features DNN's success ratio improved very little (1%-3%). That shows us the success of Random Forest in determining features.

V. CONCLUSION

ADNI data has been widely used to determine biomarkers for Alzheimer's Disease. Most of the research include MRI measurement to detect Alzheimer's disease. But, MRI measurements are not available for all patients and most of the time MRI scans are not ordered in the early stages of the disease. So creating a model that can be used with the data without MRI measurements is critical to develop low cost devices to detect Alzheimer's disease.

We created a model to obtain the most relevant features from ADNI data using baseline measurements as training and 24 months' measurements for testing purposes and then applied deep neural network to these selected features. 67% accuracy obtained using only 8 most relevant features. We believe, this is important since we eliminated MRI measurements to simplify the model and reduce the cost of the data. Proposed method can be applicable to neural network capable processors to implement fast, accurate, low-cost medical devices to detect Alzheimer's disease in early stages. Application of this method in detection of other diseases is also possible.

REFERENCES

[1] "2018 Alzheimer's disease facts and figures," Alzheimer's & Dementia, vol. 14, no. 3, pp. 367 – 429, 2018.

[2] http://adni.loni.usc.edu (ADNI Data)

[3] R. Peterson , R. Roberts, D. Knopman , B. Boeve, Y. Geda, R. Ivnik, et al. "Mild cognitive impairment: ten years later," Arch Neurol 2009; pp. 1447-1455

[4] S. Mueller, M. Weiner, L. Thal, R. Petersen, C. Jack, W. Jagust, et al. "Ways toward an early diagnosis inAlzheimer's disease: theAlzheimer's Disease Neuroimaging Initiative". Cogn. Dement 2006, pp.56–62.

[5] C. Clark, C. Davatzikos, A. Borthakur, A. Newberg, S. Leight, V. Lee. at al. "Biomarkers for early detection of Alzheimer pathology," NeuroSignals, vol. 16, no. 1, pp. 11–18, 2007.

[6] S. Minhas, A. Khanum, F. Riaz, A. Alvi, and A. Khan "A Nonparametric Approach for Mild Cognitive Impairment to AD Conversion Prediction: Results on Longitudinal Data," IEEE Journal of Biomedical & Health Informatics – 2017, vol.21, pp. 1403-1410.

[7] T. Hastie, R. Tibshirani, G. James, and D. Witten, "An Introduction to Statistical Learning: With Applications in R," Springer, CA, 2015.

[8] F. Li, A. Karpathy, J. Johnson, Standford CS Class CS231n "Convolutional Neural Networks for Visual Recognition", retrived from http://cs231n.stanford.edu

[9] V. Sze, Y. Chen, T. Yang, and J. Emer, 2017, Efficient Proccessing of Deep Neural Networks: A Tutorial and Survey, Proceedings of the IEEE, Vol. 105, pp. 2295-2329.

[10] D. Stathakis, "How many Hidden Layers and Nodes," International Journal of Remote Sensing, vol. 30, pp. 2133-. 2147, April, 2009.