Journal of Clinical Neuroscience 42 (2017) 186–192

Contents lists available at ScienceDirect

Journal of Clinical Neuroscience

journal homepage: www.elsevier.com/locate/jocn

Tools and Techniques

Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms

Füsun Er [a,*], Pınar Iscen [b], Sevki Sahin [c], Nilgun Çinar [c], Sibel Karsidag [c], Dionysis Goularas [d]

[a] Department of Biotechnology, Graduate School of Natural and Applied Sciences, Yeditepe University, Istanbul, Turkey
[b] Department of Neuroscience, Experimental Medicine and Research Institute, Istanbul University, Istanbul, Turkey
[c] Department of Neurology, Medical School, Maltepe University, Istanbul, Turkey
[d] Department of Computer Engineering, Faculty of Engineering, Yeditepe University, Istanbul, Turkey

ARTICLE INFO

Article history:
Received 29 December 2016
Accepted 6 March 2017

Keywords:
Age-related cognitive decline
Dementia, machine learning
Mild cognitive impairment

ABSTRACT

Background and aim: This study aims to examine the distinguishability of age-related cognitive decline (ARCD) from dementias based on some neurocognitive tests using machine learning.
Materials and methods: 106 subjects were divided into four groups: ARCD ($n = 30$), probable Alzheimer's disease (AD) ($n = 20$), vascular dementia (VD) ($n = 21$) and amnestic mild cognitive impairment (MCI) ($n = 35$). The following tests were applied to all subjects: The Wechsler memory scale-revised, a clock-drawing, the dual similarities, interpretation of proverbs, word fluency, the Stroop, the Boston naming (BNT), the Benton face recognition, a copying-drawings and Öktem verbal memory processes (Ö-VMPT) tests. A multilayer perceptron, a support vector machine and a classification via regression with M5-model trees were employed for classification.
Results: The pairwise classification results show that ARCD is completely separable from AD with a success rate of 100% and highly separable from MCI and VD with success rates of 95.4% and 86.30%, respectively. The neurocognitive tests with the higher merit values were Ö-VMPT recognition (ARCD vs. AD), Ö-VMPT total learning (ARCD vs. MCI) and semantic fluency, proverbs, Stroop interference and naming BNT (ARCD vs. VD).
Conclusion: The findings show that machine learning can be successfully utilized for distinguishing ARCD from dementias based on neurocognitive tests.

1. Introduction

Age-related cognitive decline (ARCD) is a non-pathological decline in memory associated with normal aging. ARCD was described by Levy in 1994 [1]. It was defined by a performance that is at least one standard deviation below age concordant values in at least one of any of the cognitive domains including learning, memory, attention, language and visuo-spatial abilities [2]. The prevalence of ARCD have been found in ratio of 20–27% in population based studies [3–5].

Mild cognitive impairment (MCI) was firstly described by Reisberg in 1982 [6]. In addition, Petersen developed this term as an intermediate stage between normal cognition and dementia. Mild cognitive impairment can be grouped as amnestic MCI (aMCI) and non-amnestic MCI (non-aMCI) based on the most affected cognitive ability. The memory in aMCI and thinking abilities in non-aMCI are decreased dramatically [3]. Cognitive deficits in MCI exceed ARCD. Moreover, MCI may progress to Alzheimer's disease (AD) [2]. Neurocognitive assessment is performed to evaluate cognitive skills of dementia patients. The test scores of a patient provide useful diagnostic information to physicians for diagnosing, treating and monitoring the disease [3]. Utilizing computational techniques including machine learning algorithms to interpret

Abbreviations: AD, Alzheimer's disease; ANOVA, analysis of variance; ARCD, age-related cognitive decline; AROC, area under the receiver-operating curve; BNT, Boston naming test; CDR, clinical dementia rating scale; IPA, International Psychogeriatric Association Working Party; MCI, mild cognitive impairment; MLP, multilayer perceptron; M5P, M5-model trees; NCT, neurocognitive test; NCTB, neurocognitive test battery; NIMH, National Institute of Mental Health; NINCDS-ADRDA, Alzheimer's Disease and Related Disorders Association; NINDS-AIREN, National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences; Ö-VMPT, Öktem verbal memory processes test; SVM, support vector machine; WAIS-R, Wechsler adult intelligence scale-revised; WMS-R, Wechsler memory scale-revised; VD, vascular dementia.

* Corresponding author at: Yeditepe University, Faculty of Engineering, 26 Agustos Campus, Kayisdagi Street 34755 Atasehir, Istanbul, Turkey.
E-mail address: fusun.er@std.yeditepe.edu.tr (F. Er).

http://dx.doi.org/10.1016/j.jocn.2017.03.021
0967-5868/

neurocognitive test (NCT) scores is an emerging area of research aiming to help physicians in diagnosis and treatment [7].

In this study, the predictive power of psychological assessment of neurocognitive abilities is evaluated in the task of distinguishing ARDC from dementias using machine learning algorithms.

2. Materials and methods

2.1. Participants

In our dementia data bank, the total number of subjects who belonged to ARCD, probable AD, MCI and vascular dementia (VD) groups was 125. 19 cases were excluded from the study due to methodological issues related with the completion of all neuropsychological tests and significant variation of age and education variables compared to average values of the groups. In total, 106 subjects (mean age = 71.97 ± 7.37 years) who were similar in age, sex and educational level were enrolled in this study.

Each subject was assigned to a clinical dementia rating scale (CDR) [8] in which a value of zero indicates no dementia and values of 0.5, 1, 2 and 3 indicate very mild, mild, moderate and severe dementia, respectively. AD and VD cases, who were classified by clinicians with a CDR value of 0.5 or 1 were selected.

Subjects were grouped as follows: 30 patients were diagnosed with ARCD (age = 70.07 ± 7.90 years, education = 10.57 ± 3.97 years, 22 female, 8 male) according to criteria described by the National Institute of Mental Health (NIMH) work group [9], 20 were diagnosed with probable AD (age = 75.45 ± 5.53 years, education = 9.75 ± 4.30 years, 11 female, 9 male) according to diagnostic criteria of National Institute of Neurological and Communicative Disorders and Stroke and by the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) [10], 21 were diagnosed with probable VD (age = 72.09 ± 6.61 years, education = 9.14 ± 3.59 years, 13 female, 8 male) according to the National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences (NINDS-AIREN) criteria [11] and 35 were diagnosed with amnestic MCI (age = 71.54 ± 7.94 - years, education = 10.34 ± 3.68 years, 18 female, 17 male) according to the criteria defined before [12].

ARCD cases were diagnosed according to the criteria of the International Psychogeriatric Association Working Party (IPA). The first criterion, the subjective impairment, is a report demonstrating that cognitive function has declined and the second one, the objective impairment, shows the difficulties in any of these cognitive tests such as memory and learning, attention and concentration, abstract thinking (problem solving, abstraction), language and visuo-spatial functioning in at least one standard deviation measurement (SD) below the age and education norms. Psychometric tests have normal values according to age and education level [13].

2.2. Neurocognitive test battery

All subjects were examined individually with a neurocognitive test battery (NCTB) that aimed to assess the functions of orientation, attention-executive functions, language, visuo-spatial skills and memory. To obtain the measures of cognitive functions of the subjects, the Turkish versions of the tests were administered. For the assessment of orientation, scores of personal and current information and the Benton judgment of line orientation test [14], were used. Attention-executive functions were examined with the following tests: digit span forward and digit span backward subscales of the Wechsler memory scale-revised (WMS-R), mental control subscale of the WMS-R, verbal fluency tests (that include naming animals for semantic fluency and K, A, S letters for phonemic fluency) [15], the similarities subtest of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) [16], interpretation of Turkish proverbs (three proverbs, with the rating of zero to three points each), the Turkish version of word fluency tests [15], a clock drawing test (a zero-to-three point scoring system was used) [17] and the interference score, incorrect response score and spontaneous correction scores of the Stroop test [18]. To assess language function, the Boston Naming Test (BNT) was performed [19]. Verbal fluency tests (both semantic and phonemic) were classified in language functions in addition to the attention-executive functions. Visuo-spatial skills were examined with the Benton face recognition test, the Benton judgment of line orientation test (visuo-spatial perception) and a copying drawings test (three simple and one complex drawings, with the rating of one point each) [14]. Both visual and verbal memory were assessed. Visual memory was examined with the visual reproduction subtest of the WMS-R including immediate recall, delayed recall and recognition scores. Verbal memory was examined with the Öktem Verbal Memory Processes (Ö-VMPT) test including scores of immediate memory, total learning, the highest learning, delayed recall and recognition [20]. WMS-R logical memory subscale is also used for verbal memory assessment.

The above tests are presented analytically in Table 2.

2.3. Data analysis

Statistical data analysis tasks are conducted using MATLAB (MATLAB and Statistics Toolbox Release 7.3, The MathWorks, Inc., Natick, MA, USA). Socio-demographic profiles among four groups were compared by a Kruskal–Wallis test [21], which is a non-parametric alternative to the one-way analysis of variance (ANOVA) [22]. The Mann–Whitney test [23] is employed to compare all possible pairs of treatment groups on test scores. All statistical tests for the features are interpreted at a significance level of 0.01.

A subset of the neurocognitive test scores was selected for each possible pair of diagnostic groups based on the nonparametric discriminatory power of them. The Benton judgment of line orientation test was not included in the analysis, as the score of this test was missing for a significant number of the cases (31 out of 106).

2.4. Machine learning algorithms

The classification via regression with M5-model trees (M5P), multilayer perceptron (MLP) and support vector machine (SVM) classifiers were constructed in order to perform the diagnosis based on the dataset with the subset of selected neurocognitive test scores. Additionally, we repeated the above procedure with the whole data set of neurocognitive test scores in order to show the importance of removing test scores in terms of success rate. The classifiers were implemented using the WEKA software [24].

A multilayer perceptron is a feed-forward artificial neural network model, which is a layered directed graph to model the relation between a set of input data to a set of output data. It is a powerful method to distinguish the data that is not linearly separable. The model is trained by a back propagation algorithm. Two important parameters are used to fine-tune the back propagation learning procedure: learning rate and momentum. The learning rate affects the converge speed to the optimum solution and momentum helps the system to avoid local minima. In this study, a learning rate of 0.3 and a momentum of 0.2 were used in the learning of MLPs.

Support vector machine is a supervised machine learning algorithm that searches an optimum hyperplane for linearly separable data. Support vector machines can be used for non-linear data