Machine learning Techniques for Diagnostic Differentiation of Mild Cognitive Impairment and Dementia

Alyssa Weakley[1], Jennifer Williams[2]

[1]Washington State University, Department of Psychology

[2]Washington State University, School of Electrical Engineering and Computer Science

Introduction

• Accurate and cost efficient classification of mild cognitive impairment (MCI) and Alzheimer's disease (AD) has benefits of personal and medical importance.

•Based on current literature, it appears that machine learning methods are viable dichotomous classifiers (e.g., Chen et al., 2010; Quintana et al., 2012). However, differentiating between MCI, AD, and neurological healthy older adults using a single model has been problematic.

The purpose of the present research is to use neuropsychological and demographic data to predict:

1. Clinical Dementia Rating (CDR) ratings
2. Clinical diagnoses (cognitively healthy, MCI, AD)

Through the implementation of four machine learning models, and a comparison statistical method:

1. C4.5 Decision tree
2. Naïve Bayes classifier (NB)
3. Artificial Neural Network (ANN)
4. Support Vector Machine (SVM)
5. Multinomial Logistic Regression

A secondary goal of the project is to determine the fewest number of attributes (e.g., neuropsychological tests) required to reliably diagnose individuals.

Method

Participants

• For participant characteristics see Tables 1 and 2. Clinical diagnoses were consistent with Petersen and colleagues criteria, NIA-Alzheimer's Association workgroup, and NINCDS-ADRDA criteria. CDR ratings were provided by a certified rater.

Table 1: Demographic and feature selected neuropsychological data of participants classified using clinical diagnosis

	Control (n = 161)		MCI (n = 97)		AD (n = 53)		*p*
Variable or test	Mean	SD	Mean	SD	Mean	SD	
Age (years)	71.14	8.47	71.96	9.43	75.70[ab]	8.26	.005
Education (years)	16.28	2.83	15.52	2.93	16.11	2.83	.04
% Female	74	-	57[a]	-	45[a]	-	<.001
TICS	34.85	2.27	32.48[a]	2.99	24.36[ab]	5.84	<.001
Trails B	79.45	27.37	123.8[a]	58.4	235.25[ab]	88.03	<.001
Clox 1	12.92	2.08	11.84[a]	2.36	8.98[ab]	3.47	<.001
Design Fluency – Open dots	10.06	2.87	8.27[a]	2.76	4.41[ab]	2.47	<.001
Design Fluency – Switching	7.63	2.56	5.88[a]	2.65	2.18[ab]	2.00	<.001
Functional Status z-score	-.31	.60	.10[a]	.83	1.06[ab]	1.82	<.001
Depression z-score	-.24	.83	.16[a]	.94	.18[a]	1.13	.001
Delayed Verbal Memory z-score	.40	.59	-.19[a]	1.02	-1.24[ab]	1.29	<.001
Visual Memory total correct z-score	.60	.67	-.38[a]	.80	-1.3[ab]	.72	<.001

Note: Scores are raw scores unless otherwise listed. AD = Alzheimer's disease; MCI = mild cognitive impairment; TICS = Telephone Interview for Cognitive Status. L-N = Letter Number Sequencing.
[a]differed significantly from control group; [b]differed significantly from MCI group.

Table 2: Demographic and feature selected neuropsychological data of participants classified using clinical diagnosis

	0 (n = 154)		0.5 (n = 93)		1.0 (n = 25)		*p*
Variable or test	Mean	SD	Mean	SD	Mean	SD	
Age (years)	70.06	9.39	74.04	7.97	71.95	9.34	.08
Education (years)	16.13	2.86	15.53	2.89	16.12	2.89	.26
% Female	71	-	58[a]	-	68[a]	-	<.001
TICS	34.66	3.36	32.40[a]	3.36	23.92[ab]	2.38	<.001
Trails A	34.17	11.51	42.50[a]	3.36	81.76[ab]	58.49	<.001
Letter Fluency	41.17	11.83	35.90[a]	9.69	22.67[ab]	8.95	<.001
Category Fluency	41.15	8.12	35.09[a]	9.69	22.09[ab]	8.96	<.001
Category Switch	13.81	2.71	11.24[a]	3.27	6.18[ab]	3.03	<.001
Design Fluency – Solid dots	8.95	3.01	7.16[a]	3.17	4.18[ab]	2.28	<.001
Functional Status z-score	-.25	.69	.17[a]	.85	1.72[ab]	1.67	<.001
Immediate Memory z-score	.31	.64	-.08[a]	1.06	-1.09[ab]	1.12	<.001
Visual Delayed Memory z-score	.47	.77	-.18[a]	.91	-1.31[ab]	.87	<.001
Visual Memory z-score	.49	.78	-.21[a]	.89	-1.32[ab]	.71	<.001

Note: Scores are raw scores unless otherwise listed. AD = Alzheimer's disease; MCI = mild cognitive impairment; F = female; M = male; TICS = Telephone Interview for Cognitive Status; L-N = Letter Number Sequencing.
[a]differed significantly from CDR = 0; [b]differed significantly from CDR = 0.5.

Neuropsychological and Demographic Measures

• Demographic variables: age, education, and gender
• Mental status: Telephone Interview of Cognitive Status (TICS)
•Depression: Geriatric Depression Scale (GDS)
•Functional Statues
•Neuropsychological Performance:
  • Attention/speeded processing (Symbol Digit Modalities Test, Trails A)
  • Verbal learning and memory (RAVLT, Memory Assessment Scales)
  • Visual learning and memory (7/24, Brief Visual Memory Test)
  • Executive functioning (Trails B, Clox 1, Design Fluency of D-KEFS)
  • Working memory (WAIS-III Letter-Number sequencing)
  • Verbal fluency (verbal fluency subtest from the D-KEFS)
  • Confrontation naming (BNT)
  • Word knowledge (Shipley Institute of Living Scale)

Results

Feature Selection

•CDR-FS
  •Age
  •Education
  •Gender
  •Functional Status
  •Visual memory total correct
  •Trails A
  •Letter Fluency
  •Category Fluency
  •Switch Fluency
  •Design Fluency – Solid Dots
  •Short delay verbal memory
  • Long delay visual memory

•Clinical Diagnosis-FS
  •Age
  •Education
  •Gender
  •Functional Status
  •Visual memory total correct
  •TICS
  •Trails B
  • Clox1
  •Design Fluency – Open Dots
  • Design Fluency – Switching
  • Depression
  • Long delay verbal memory

Table 3: Results from supervised model for original, reduced, and feature selection datasets. This table shows the results of modeling for the clinical diagnosis class with and without CDR total

		CD	CD-r	CD-r_fs	CD_no-cdr	CD-r_no-cdr	CD-r_no-cdr_fs
NB	missing	**80.7%**	**83.9%**	**92%**	76.2%	78.5%	**83.3%**
	not missing	**85.2%**	**85.2%**	**93.3%**	77.8%	77.5%	**81.7%**
DT	missing	74.3%	76.2%	**81.3%**	65.9%	68.5%	74.3%
	not missing	**88.7%**	**90.1%**	**91%**	79.7%	71.1%	78.5%
SVM	missing	79.1%	66.2%	65.6%	70.7%	62.7%	60.1%
	not missing	**86.1%**	79.1%	79.4%	79.4%	68.5%	69.4%
NN	not missing	**89.7%**	**90.4%**	**91.6%**	**80.3%**	78.1%	**82%**
Logistic Regression	not missing	-	-	-	-	-	**86.5%**

Note: missing = missing attribute values, not missing = missing values were replaced with average

Table 4: Results from supervised model for original, reduced, and feature selection datasets, CDR class

		CDR	CDR-r	CDR-r_fs
NB	missing	73.9%	75%	**80.1%**
	not missing	**81.6%**	75.75	79.4%
DT	missing	70.2%	65.4%	74.3%
	not missing	**81.6%**	68.4%	73.9%
SVM	missing	64.3%	64.3%	65.8%
	not missing	**80.9%**	69.2%	70.6%
NN	not missing	76.9%	75%	77.2%
Logistic Regression	not missing	-	-	77.9%

Note: missing = missing attribute values, not missing = missing values were replaced with average

Table 5: Results from semi-supervised model for original, reduced, and feature selection datasets.

		ss-CDR	ss-CDR-r	ss-CDR-r_fs
NB	missing	78.7%	**80.3%**	**82.3%**
	not missing	**88.8%**	**81.2%**	**83.4%**
DT	missing	68.3%	73.9%	**80.3%**
	not missing	**81.2%**	73%	71.9%
SVM	missing	69.7%	73.9%	76.7%
	not missing	**86%**	74.2%	77.8%
NN	not missing	**80.6%**	**80.6%**	**80.6%**

Note: missing = missing attribute values, not missing = missing values were replaced with average

Conclusions

•We hypothesized that NB and SVM would provide the greatest accuracy. Based on the results, NB achieved the highest accuracy in all cases. However, our prediction that SVM would also provide a high accuracy rate, was not supported.
•Improved accuracy over a traditional statistical method was observed for CDR classification but not for clinical diagnosis.
•We explored the use of feature selection to reduce the number of demographic and neuropsychological data needed to make an accurate classification. We were able to determine which tests were critical when making classification for CDR scores and clinical diagnoses.
•The experiments reported in this paper indicate that artificial intelligence techniques can be used to automate aspects of clinical diagnosis and can provide meaningful insights into which attributes are the most valuable for this diagnosis.