Published in final edited form as:

Soc Psychiatry Psychiatr Epidemiol. 2018 January ; 53(1): 77–86. doi:10.1007/s00127-017-1410-0.

Application of machine learning methods for diagnosis of dementia based on the 10/66 battery of cognitive function tests in south India

S R Bhagya Shree [Research Scholar],

PET Research Centre, PES College of Engineering, Mandya, India, 9900110944

Dr H S Sheshadri [Dean],

PET Research centre, Prof. Department of Electronics & Communication Engineering, PES College of Engineering, Mandya, India

Mr Kiran Nagaraj,

CSI Holdsworth Memorial Hospital, PO Box 28, Mandimohalla, Mysore, India

Prof Martin Prince [Professor of Epidemiological Psychiatry],

Institute of Psychiatry, Kings College, London, UK

Prof Caroline HD Fall [Professor of International Paediatric Endocrinology], and

MRC Life course Epidemiology Unit. University of Southampton, UK

Dr Murali Krishna* [Wellcome DBT Early Career Fellow and Consultant Psychiatrist]

CSI Holdsworth Memorial Hospital, PO Box 28, Mandimohalla, Mysore, India

Abstract

Back ground—There is limited data on the use of Machine learning methods for automating clinical aspects of dementia in low and middle income country (LMIC) settings including India. A culture and education fair battery of cognitive tests was developed, validated and normed for use in LMICs including south India by the 10/66 Dementia Research Group. We explored the machine learning algorithms to determine if the analysis of neuropsychological data from the 10/66 battery of cognitive tests can be automated for the diagnosis of dementia in south India.

Methods—The data sets for 466 men and women aged 55- 80 yrs were obtained from the on-going Mysore Studies of Natal effect of Health and Ageing (MYNAH), in south India. This includes subject demographics, performance on the 10/66 cognitive function tests, diagnosis of mental disorders and population based normative data for the 10/66 battery of cognitive function tests. We examined the diagnostic properties of the battery of cognitive tests and derived an equation to enhance the accuracy of diagnosis of dementia. Machine learning techniques were applied to the data set.

Results—Of 466 subjects, 27 had 10/66 diagnosis dementia. 19 of them were correctly identified as having dementia by Jrip classification with 100% accuracy.

(Corresponding author) muralidoc@gmail.com Phone: 0091991658550 Fax 00918214007000at.

Conflict of Interest

None of the authors have any conflict of interest to declare.

Conclusions—This pilot exploratory study indicates that machine learning methods can help to identify community dwelling older adults with 10/66 criterion diagnosis of dementia with good accuracy in a LMIC setting like India. This should reduce the duration of the diagnostic assessment and make the process easier and quicker for both the clinicians and patients.

Keywords

machine learning; Jrip; dementia diagnosis; 10/66 cognitive function tests; Low and middle income setting

Introduction

Neurocognitive disorders are a major cause of disability and mortality in late life and are associated with high costs for health systems and society particularly in low and middle income countries (LMIC) like India [1] [2]. Population based studies in India report prevalence rates of 7.5% and 10.6% for dementia in those aged above 60 yrs in urban and rural areas respectively [2][3]. The proportion of persons with dementia in India is expected to increase two-fold by 2030 because of the steady growth in the older population and stable increments in life expectancy [1] [2]. The Global Burden of Disease report identifies dementia as one of the main causes of disability and this has a disproportionate impact on capacity for independent living in later life [4]. Although neurocognitive disorders are the second highest source of burden after tropical diseases, research into improving the diagnosis in LMIC setting remains minimal [1][2].

Instruments and batteries of neuropsychological tests for screening and diagnosis of dementia are widely available for use in high income setting. They include: the Mini Mental State Examination (MMSE), Blessed Orientation Memory and Concentration test (BOMC), Montreal Cognitive Assessment (MOCA), The Eight Item Interview to differentiate ageing from dementia (AD8) and General Practitioner Assessment of cognition (GP CoG). They are of limited use in LMIC settings where literacy levels are low and the cultural context is radically different to the west [4]. Therefore, a culture and education fair battery of cognitive tests was developed, validated and normed for use in LMICs including south India by the 10/66 Dementia Research Group. This is suitable for use in people with little or no education [5]. The 10/66 battery of cognitive tests comprises: the Community Screening Instrument for Dementia (CSID) incorporating the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) animal naming verbal fluency task, the modified CERAD 10 word list learning task with delayed recall, and an informant interview for evidence of cognitive and functional decline [5][6]. In the 10/66 pilot studies, the CSID, informant interview and the modified CERAD 10 word list-learning task were independently able to predict the diagnosis of dementia [5].

Accurate and early diagnosis of dementia has benefits of personal and medical importance [7]. Diagnosis of dementia in clinical settings is time intensive and requires multiple sources of information (e.g. neuro psychological assessments, laboratory reports and reliable informant reports). These data are assembled to create a cohesive picture of the individual's impairment to aid clinical diagnosis, where efficiency and accuracy are governed by the

practitioner's level of expertise. Limited diagnostic capacity and the monetary expense of a medical diagnosis are major concerns in LMIC settings [4]. There is an urgent need to develop reliable alternatives to traditional medical diagnosis [8].

Studies from high income settings report that artificial intelligence techniques can be used to automate aspects of clinical diagnosis in individuals with cognitive impairment [9]. Data mining is the process of sorting and categorising data into various types, forms or any other distinct class. The most commonly employed classifiers for data mining and their key characteristics are listed in table one.

Data mining is an emerging field of high importance for providing diagnosis and evaluating prognosis of dementia. Machine learning algorithms have been employed by Williams and colleagues for diagnostically differentiating mild cognitive impairment and dementia using the Clinical Dementia Rating [10]. Schmitter-Edgecombe and colleagues evaluated machine learning models and report that the application of Naïve bayes and neural networks provided higher accuracy than SVM (Support Vector Machine) and decision tree methods. They report 67% sensitivity for diagnosis of mild cognitive impairment (MCI) through evaluation of multiple memory processes and exploration of day to day activities [11]. Shankle and colleagues used machine learning algorithms namely C4.5 and C4.5 rules for the diagnosis of dementia based on the Blessed Orientation Memory and Concentration (BOMC) test. The probability of correct diagnosis using C4.5 & C4.5 rules was 85.5% & 85.9% respectively [12]. Application of machine learning algorithms like Naïve Bayes, C4.5 and IB1 (Instance-based learning version 1) for the diagnosis of Alzheimer's disease based on Functional Activities Questionnaire (FAQ) and BOMC tests by Datta and colleagues improved classification accuracy by 15-20% [13]. Joshi and colleagues demonstrated the accuracy of staging dementia could be optimised by combining results of FAQ and MMSE by employing machine learning and neural networks [14].

There are limited data on the use of artificial intelligence methods for automating clinical aspects of dementia in LMIC settings including India. We are not aware of any studies that have examined the utilisation of the 10/66 data by artificial intelligence methods for accurate diagnoses of dementia. Therefore, we explored machine learning algorithms to determine if the analysis of neuropsychological data from the 10/66 battery of cognitive tests, along with demographic and population based normative data, can be automated for the diagnosis of dementia in south India.

Aims

Arriving at a 10/66 diagnosis of dementia requires the subject to undergo a combination of mental health, neurological and cognitive assessments listed in table one and takes 2-3 hours. In addition, a reliable informant must also be interviewed using a structured questionnaire for evidence of functional impairment resulting from cognitive decline. We aimed to apply machine learning techniques to identify the minimum number of cognitive function tests required for diagnosis of dementia and derive an algorithm to improve the diagnostic accuracy. This should reduce the duration of the diagnostic assessment and make the process easier and quicker for both the clinicians and patients.

Methods

a Setting

This study was carried out at the Epidemiology Research Unit, CSI Holdsworth Memorial Hospital (HMH), Mysore, South India. The study was approved by the HMH Ethics and Research Committee.

b Sample description

Data for this study were obtained from the on-going Mysore Studies of Natal effects on Health and Ageing (MYNAH), in south India [15]. They are members of the Mysore Birth Records Cohort, aged 55-80 yrs, and have undergone a comprehensive assessment for cognitive function, mental health and cardiometabolic disorders. This study commenced in March 2013 and data for the first 466 men and women, who participated before September 2014 were included. Characteristics of the participants are provided in table two.

Assessments for the 10/66 diagnosis of dementia include the 10/66 battery of cognitive tests, a Geriatric Mental State examination and a brief structured neurological examination; Community Screening Instrument for Dementia (CSID) informant interview for evidence of cognitive and functional decline and a Neuropsychiatric Inventory for behavioural and psychological symptoms of dementia. [4] (Refer table three). Dementia is defined by a score above a cut-off point of predicted probability of DSMIV Dementia Syndrome from the logistic regression equation of the 10/66 dementia diagnostic algorithm [16, 17].

c Cognitive function tests

Cognitive functioning as a continuous measure was obtained by administering the Kannada (local language) version of the 10/66 cognitive assessment battery and was validated for use in this population [18]. This is drawn principally from the Community Screening Instrument for Dementia (CSID) developed by the Ibadan-Indianapolis study group [19], specifically for use in cross-cultural research, and in low education settings, and from the Consortium to Establish a Registry for Alzheimer's Disease CERAD [24]. This battery comprises:

- i) Global Cognitive Function measured by administering the Community Screening Instrument for Dementia (CSI 'D') to the subjects [19]. This includes a 32 item cognitive test assessing orientation, comprehension, memory, naming and language expression, which generates a global cognitive score (CSI'D' COGSCORE). The CSI'D' COGSCORE score distributions among those with dementia and controls, and the degree of discrimination provided was remarkably consistent across different cultural settings [19][24–27].
- ii) Verbal fluency (VF) measured by the animal naming verbal fluency task from the CERAD [19, 24]. After a brief practice, naming items from another category (clothing), participants are encouraged to name as many different animals as they can in the space of one minute. The instructions read out to the participant stipulate: 'think of any kinds of animal in the air, on land, in the water, in the forest, all the different animals'. If the participant stops before the allotted time

has elapsed they are encouraged to continue. They score one point for each valid name.

- iii) Memory is measured by the modified Word List Memory and Recall test. Word List Recall (WLR), a measure of delayed recall, has been reported to be of particular value in distinguishing early dementia from normal aging [28]. The word list is taken from the adapted CERAD ten word list learning task used in the Indo-US Ballabgarh dementia study [29]. Six words- butter, arm, letter, queen, ticket and grass were taken from the original CERAD battery English language list [30]. Pole, shore, cabin, and engine were replaced with corner, stone, book and stick, which were deemed more cross-culturally applicable [5]. In the learning phase, the list is read out to the participant from a green card, who is then asked to recall straight away the words that they remember. Approximately five minutes later, after a series of unrelated CSI'D' questions (name registration, object naming, object function and repetition) the participant is again asked to recall the 10 words with prompting that they were read from a green card, giving a WLR score out of 10.

d Other data

Demographic data, including age, education and gender; performance in the three cognitive tests (listed above) that were administered by a trained clinical psychologist; and population based normative data for the 10/66 battery of cognitive tests for urban south India (appendix one).

e Data evaluation

Stage 1: For the pilot study, we initially applied several diagnostic algorithms namely Jrip, Naïve Bayes, Random Forest and J48 on a simulated dataset of 250 men and women for making a diagnosis of dementia. Naïve Bayes, Jrip and Random Forest algorithms performed better than J48 for making a diagnosis of dementia [31].

We then applied Naïve Bayes and J48 algorithms to primary data from the 466 study participants from the MYNAH study. SMOTE (Synthetic Minority Oversampling Technique), a pre-processing technique was applied to reduce the imbalance in the data. Application of SMOTE significantly expanded the data and therefore was deemed unnecessary. The Knowledge explorer interface of Weka tool was used to classify the pre-processed data into groups with and without dementia. Naïve Bayes performed better than J48 for making a diagnosis of dementia [32]. Subsequently we evaluated Jrip, Random forest and Naïve Bayes classifiers using explorer, Knowledge flow & Java API (Application Programming in Interface). The results of explorer, knowledge flow and Java API were similar and confirmed that results would be the same irrespective of the mode of model evaluation. Irrespective of number of subjects, methods of model evaluation and classification techniques used, Jrip consistently performed better in our analyses (results provided in appendix two).[33].

Stage 2: We compared the results from stage 1 with that of the 10/66 diagnosis of dementia and examined the level of agreement. Circumstances leading to any disagreement are provided in the results section.

Stage 3: To minimise the level of disagreement, we proposed a data processing pathway for classification (Figure one) comprising a series of sequential processes.

Diagnostic cut offs were provided for each of the three cognitive tests at the mean, 0.5, 1, 1.5 and 2 standard deviations lower than the population mean for urban south Indians provided by the 10/66 Dementia Research Group [16]. We considered diagnostic cut-offs for cognitive tests individually and in varying combinations. The following equations for diagnosis of dementia were examined:

- a) A score lower than the specified cut off for all the three cognitive tests
- b) A score of any one of three cognitive tests lower than the specified cut-off
- c) A score of any two of the three cognitive tests lower than the specified cut-offs.
- d) A score lower than 0.5 SD for either CSID COGSCORE or Verbal fluency and a WLR score lower than 0.5 SD.
- e) A score lower than the mean for either CSID COGSCORE or Verbal fluency and WLR score lower than 0.5 SD.

Stage 4: The true positive (diseased and test positive), false positive (not diseased and test positive), false negative (diseased and test negative) and true negative (not diseased and test negative) cases were recorded. The sensitivity (percentage of patients with the disease that receive a positive result), specificity (percentage of patients without the disease that receive a negative result), positive predictive value (percentage chance that a positive test result is a true positive) and negative predictive value (percentage chance that a negative test result is a true negative) were calculated by considering the results obtained from 10/66 diagnostic algorithm as the gold standard. The confidence interval (CI) with $\alpha=0.05$ was calculated by using the following formulae

$CI = \text{Estimate} \pm 1.96 \times SE$ (where $SE = \text{Proportion} \times (1 - \text{Proportion}) / \text{Denominator of the proportion}$).

Estimate for sensitivity = True Positive / True Positive + False Negative.

Estimate for Specificity = True Negative / True Negative + False Positive.

Estimate for Positive Predictive value = True Positive / True Positive + False Positive.

Estimate for Negative Predictive Value = True Negative / True Negative + False Negative [34]

After the calculation, the results were tabulated (table four). This resulted in thirty two different combinations of cognitive function test with varying levels of diagnostic cut offs. From the table three it is apparent that providing weightage to WLR would improve the diagnostic accuracy. Hence we defined dementia as scoring lower than the population mean

for either VF or CSID COGSCORE with a WLR score 1 standard deviation lower the population mean.

Dementia = Score lower than the mean for VF or CSI'D' COGSCORE + Score lower than 1 SD for WLR.

For the diagnosis of dementia, the above equation provided the optimal kappa ($k=0.81$) value with satisfactory sensitivity ($=0.7$), specificity ($=1$), positive predictive value ($=1$) and negative predictive value ($=0.98$) was chosen for further processing.

Stage 5 Preparing the data: The data collected was represented in the form of spread sheet. The native storage format of Weka is ARFF (Attribution Relation File Format). Therefore, the spread sheet data was initially transformed to a CSV (Comma Separated Value) format and further to an Attribution Relation File Format format for preprocessing [35].

Stage 6 Preprocessing: The data from stage 5 was preprocessed for model evaluation and feature selection. Application of preprocessing techniques, namely imbalance reduction, discretisation and randomisation were not required [32]. Features were selected using the best search technique of the Wrapper method. Wrapper method is the recursive feature elimination algorithm. Wrapper methods consider the selection of a set of features as a search problem, where different combinations are prepared, evaluated and compared to other combinations. [36] The 54 attributes derived from cognitive tests were reduced to 8 attributes. By feature selection the total number of features was substantially reduced and application of classification after the removal of unwanted features appeared to improve the accuracy of classification. But, the reduced features included COGSCORE, VF and WLR and evidently to calculate the CSI 'D' COGSCORE all the variables from the datasets were required. This confirmed that all the items mentioned in the cognitive tests were equally important and required to derive the diagnosis of dementia.

Stage 7 Classification :

The data were classified using Jrip, which uses an IF – THEN classification. The classification was conducted before and after feature selection and applied to the dataset with and without reduced features. The model was evaluated using training and cross validation by applying Jrip classification.

Results

1. Of the 466, 27 had the 10/66 diagnosis of dementia. Of these, 19 were identified correctly as having dementia by the Jrip algorithm. 8 subjects with the 10/66 diagnosis of dementia were classified as having 'no dementia' (false negatives). There were no false positives. The circumstances for disagreement were one or more of the following

- (i) The 10/66 diagnosis of dementia was based on the eight tests whereas the machine learning diagnosis was based on the subject's performance on cognitive tests alone.
- (ii) Two subjects had major learning disability.

- (iii) In six subjects the cognitive impairment was secondary to major depressive disorders (pseudo dementia syndrome).

2. The performance measure after applying Jrip classification for dementia diagnosis is provided in table 5. Classification accuracy, Precision, Recall and F-score are the basic measures of classification. Classification accuracy is the percentage of test set tuples that are correctly classified. Precision is the percentage of retrieved documents that are relevant to the query (correct responses). Recall is the percentage of retrieved documents that are relevant to the query and were retrieved. The accuracy of Jrip classifier is 100% before and after the feature selection for both the training set and cross validation.

3. The feature selection confirmed that all items mentioned in the 10/66 battery cognitive of tests are equally important and required for diagnosis of dementia in our setting.

Discussion

Predictive data mining for medical diagnosis have been shown to significantly improve the quality of clinical decisions and eliminate unwanted biases, errors and excessive medical costs which affect the quality of service provided to patients [37, 38]. Our results from this pilot study indicate that machine learning methods can help identify older adults with 10/66 criterion diagnosis of dementia with good accuracy in a LMIC setting like India. We have demonstrated that this can be achieved by processing a dataset that combines participants' demographic data, performance on the 10/66 battery of cognitive tests, and population based normative value for these tests provided by the 10/66 Dementia Research Group for urban south Indians.

Our machine learning methods were able to separate community dwelling older adults into those with dementia and those without, making way for early diagnosis and providing an opportunity for early intervention. By using the proposed system the time taken for diagnosis of dementia is substantially reduced. More importantly non-specialists can be successfully trained to administer the 10/66 battery of cognitive tests. Therefore the development of machine learning methods for diagnosis of dementia in the community is a promising way of helping to reduce the mental health gap (mhGap), the gap between what is urgently needed and what is available to reduce the burden for neurocognitive disorders in LMICs [8].

Future work entails developing a guideline for machine learning methods that can be applied by clinicians and non-specialists for diagnosing dementia. Future work should also examine whether machine learning methods in conjunction with the other clinical data can be applied to create simple statements to help classify the dementia status of patients. In addition to these statements, prototypes learnt by machine learning methods can aid in better understanding the characteristics of each class. The proportion of those with the diagnosis of dementia in our dataset was small. Therefore we need to test our machine learning models in a larger sample, in different settings and study the pattern of predictor values for dementia including its subtypes (e.g. Alzheimer's disease, Vascular dementia etc). We will also examine if enabling the machine learning methods to differentiate depression with cognitive impairment (pseudo-dementia syndromes) and dementia will increase the accuracy further.

This study has several strengths. The datasets were obtained from community dwelling older adults as opposed to using simulated data sets and were pre-processed before submission to machine learning methods. They were complete for all the 466 subjects. The diagnostic properties for each of the cognitive function tests at varying cut-offs in relation to the population mean values were examined. The diagnosis of dementia from the rule based approach was compared against the 10/66 diagnosis of dementia (gold standard). The advantages of the rule based approach (expert system) is that the knowledge base can be updated, extended and include a large amount of information.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This study was partly funded by the Wellcome DBT India Alliance as an Early Career Research Fellowship to Dr Murali Krishna. The grant providers did not influence the design, conduct of the study and interpretation of the findings. Our sincere thanks to the participants and their families for taking part in this study. We thank Saroja A, Santhosh Nagaraj, Ramya R and Bhavya Hegde, CSI Holdsworth Memorial Hospital for assistance with recruitment and assessments this study. We thank Dr L Basavaraj, Principal, ATME College of Engineering, Mysore for the support. We thank Mr Kiran K Nagaraj, CSI Holdsworth Memorial Hospital and MS Patsy Coakley, MRC Life course Epidemiology Unit for their support with data management.

Abbreviations

CSI'D'	Community Screening Instrument for Dementia
CSID'T'	Community Screening Instrument for Dementia-Informant Interview
COAD	Chronic Obstructive Airway Disease
VF	Verbal Fluency
WLR	Word List Recall
CERAD	The Consortium to Establish a Registry for Alzheimer's Disease
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
LMIC	Low and Middle Income Country
MMSE	Mini Mental State Examination
BOMC	Blessed Orientation Memory and Concentration test
MOCA	Montreal Cognitive Assessment
AD8	Ascertain Dementia 8

GP CoG	General Practitioner Assessment of cognition
IB1	instance-based learning version 1
ARFF	Attribute-Relation File Format
CSV	Comma Separated Variable

References

- [1]. <http://www.alz.co.uk/research/files/WorldAlzheimerReport.pdf> (Link to World Alzheimer's report 2009).
- [2]. Shaji KS, Jotheeswaran AT, Girish N, Srikala BharathAmit DiasMeera PattabiramanMathew Varghese. The Dementia India Report: 2010. Prevalence, impact, costs and services for Dementia. Alzheimer's and Related Disorders Society India; New Delhi: 2010.
- [3]. Prince Martin J. The 10/66 dementia research group - 10 years on. Indian Journal of psychiatry. 2009; 51(Suppl. 1):S8–S15. [PubMed: 21416024]
- [4]. Maestre Gladys E. Assessing Dementia in Resource-Poor Regions. Current Neurology and Neuroscience Reports. 2012; 12(5):511–519. [PubMed: 22864986]
- [5]. Prince M, Acosta D, Chiu H, Scazufca M, et al. Dementia diagnosis in developing countries: a cross-cultural validation study. Lancet. 2003; 361(9361):909–917. [PubMed: 12648969]
- [6]. Prince M, Ferri CP, Acosta D, Albanese E, et al. The protocols for the 10/66 dementia research group population-based research programme. BMC Public Health. 2007; 7:165. [PubMed: 17659078]
- [7]. Prince M, Bryce R, Ferri C, editorsAlzheimer's Disease International. World Alzheimer Report London: The benefits of early diagnosis and intervention2011.
- [8]. Mental Health Gap Action Programme. scaling up care for mental, neurological and substance use disorders. Geneva: World Health Organization; 2008.
- [9]. Han JiaweiKamber MichelinePie Jian. Data Mining : Concepts and Techniques. Third edition. Elsevier; 2012.
- [10]. Williams Jennifer A, Weakley AlyssaCook Diane J, Schmitter-Edgecombe Maureen. Machine learning techniques for diagnostic differentiation of mild cognitive impairment and dementia. Expanding the Boundaries of Health Informatics Using Artificial Intelligence; Workshops at the Association for the Advancement of Artificial Intelligence Conference; 2013. 71–76.
- [11]. Schmitter-Edgecombe M, Parsey C, Cook DJ. Cognitive correlates of functional performance in older adults:comparison of self report, direct observation and performance-based measures. Journal of the international neuropsychological society. 2011; 17:853–864. [PubMed: 21729400]
- [12]. Shankle William RodmanDatta P, Dillen courtPazzani M. Improving Dementia Screening Tests with Machine learning methods. Alzheimer's Research. 1996; 2(714):95–99.
- [13]. Datta P, Shankle William RodmanPazzani M. Applying machine learning to an Alzheimer's database. Conf. proc. AAAI Symposium; 1996.
- [14]. Joshi SandhyaDeepa Shenoy P, Venugopal KR, Patnaik LM. Evaluation of different stages of dementia employing neuropsychological and machine learning techniques. First International Conference on Advanced Computing; 2009. 154–160.
- [15]. Krishna MuraliKalyanaraman KumaranVeena SR, Krishanveni GV, Karat SC, Cox VanessaCoakley PatsyNagaraj KiranStein ClaudiaPaul BDR, Prince Martin, et al. Cohort Profile: The 1934–66 Mysore Birth Records Cohort in South India. International Journal of Epidemiology. 2015; 44(6):1833–1841. [PubMed: 26445965]
- [16]. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Primary Care Version. 4th edition. Washington, DC: American Psychiatric Association; 1994.
- [17]. Sosa Ana LuisaAlbanese EmilianoPrince MartinAcosta DaisyFerri Cleusa P, Guerra MariellaHuang YueqinJacob KS, de Rodriguez Juan LlibreSalas AquilesYang Fang, et al. Population normative data for the 10/66 Dementia Research Group cognitive test battery from

- Latin America, India and China: a cross-sectional survey. *BMC neurology*. 2009; 9:1–11. [PubMed: 19126193]
- [18]. Krishna M, Beulah E, Jones S, Sundarachari R, A S, Kumaran K, Karat SC, Copeland JR, Prince M, Fall C. Cognitive function and disability in late life: an ecological validation of the 10/66 battery of cognitive tests among community-dwelling older adults in South India. *Int J Geriatr Psychiatry*. 2015; doi: 10.1002/gps.4404
- [19]. Hall KS, Gao S, Emsley CL, Ogunniyi AO, et al. Community screening interview for dementia (CSI D); Performance in five disparate study sites. *Int J Geriatr Psychiatry*. 2000; 15(6):521–531. [PubMed: 10861918]
- [20]. Copeland JRM, Dewey ME, Griffith-Jones HM. A computerised psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGE CAT. *Psychological Medicine*. 1986; 16:89–99. [PubMed: 3515380]
- [21]. Dewey ME, Copeland JR. Diagnosis of dementia from the history and aetiology schedule. *Int J Geriatr Psychiatry*. 2001; 16:912–917. [PubMed: 11571773]
- [22]. Broe GA, Akhtar AJ, Andrews GR, Caird FI, Gilmore AJ, McLennan WJ. Neurological disorders in the elderly at home. *J Neurol Neurosurg Psychiatry*. 1976; 39:361–366.
- [23]. Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, Lopez OL, DeKosky ST. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clinical Neuroscience*. 2000; 12:233–239.
- [24]. Morris JC, Heyman A, Mohs RC, Hughes JP, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD), Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*. 1989; 39(9):1159–1165. [PubMed: 2771064]
- [25]. henderie Hugh C, Hall Kathleen S, Pillay NeelanRodgers DonaldPrince CarolNorton JamesBrittain HarryNath AvindraBlue ArthurKaufert JoeShelton Paul, et al. Alzheimer's disease is rare in Cree. *Int Psychogeriatr*. 1993; 5(1):5–14. [PubMed: 8499574]
- [26]. Hendrie HC, Osuntokun BO, Hall KS, Ogunniyi AO, et al. Prevalence of Alzheimer's disease and dementia in two communities :Nigerian Africans and African Americans. *The American Journal of Psychiatry*. 1995; 152(10):1485–1492. [PubMed: 7573588]
- [27]. Unverzagt FW, Morgan OS, Thesiger CH, Eldemire DA, et al. Clinical utility of CERAD neuropsychological battery in elderly Jamaicans. *J Int Neuropsychol Soc*. 1999; 5(3):255–259. [PubMed: 10217925]
- [28]. Welsh K, Butters N, Hughes J, Mohs R, et al. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Archives of Neurology*. 1991; 48(3):278–281. [PubMed: 2001185]
- [29]. Ganguli M, Chandra V, Gilby JE, Ratcliff G, et al. Cognitive test performance in a community based non-demented elderly sample in rural India: the Indo-U.S. Cross-National Dementia Epidemiology Study. *Int Psychogeriatr*. 1996; 8(4):507–524. [PubMed: 9147167]
- [30]. Guruge O, Unverzagt FW, Osuntokun BO, Hendrie HC, et al. The CERAD Neuropsychological Test Battery: norms from a Yoruba-speaking Nigerian sample. *West Afr J Med*. 1995; 14(1):29–33. [PubMed: 7626529]
- [31]. Bhagya shree SR, Sheshadri HS. An initial investigation in the diagnosis of Alzheimer's disease using various classification techniques. *IEEE-ICCIC 2014*; 1–5.
- [32]. Bhagya shree SR, Sheshadri HS. An approach to preprocess data in the diagnosis of Alzheimer's Disease. *IEEE- CCIOT 2014*; 2014. 135–139.
- [33]. Bhagya shree SR, Sheshadri HS, krishna Murali. Diagnosis of Alzheimer's disease employing Neuropsychological and classification techniques. *IEEE- ICST 2015*; 2015.
- [34]. Fenn Buderer Nancy M, MS. Statistical Methodology : I . Incorporating the Prevalence of Disease into the Sample Size Calculation for Sensitivity and Specificity. *ACADEMIC EMERGENCY MEDICINE*. 1996 Sep; 3(9):895–900. [PubMed: 8870764]
- [35]. Soman KP, Diwakar ShyamAjay V. Insight into data mining theory and practice. New Delhi: PHI learning Pvt. Ltd;
- [36]. Kohavi RonJohn George H. Wrappers for feature subset Selection. *Artificial Intelligence*. 1997; 97(1–2):273–324.

- [37]. Herron Patrick. Machine Learning for Medical Decision Support: Evaluating Diagnostic Performance of Machine Learning Classification Algorithms. 2004
- [38]. Jain Anju. Machine Learning Techniques for Medical Diagnosis A REVIEW. Conf. Proc. ICSTM 2015; 2015. 2449–2459.

Fig one. Data evaluation flow diagram

Table one
Key characteristics of the data classifiers

No	Name of the Classifier	Key characteristics
1.	Naive Bayseian classifier	is a selective classifier which calculates the set of probabilities by counting the frequency and combination of values in a given data set. It assumes that all variables which contribute towards classification are mutually independent. Naïve Bayesian classifier is based on Bayes theorem and theorem of total probabilities. [9]
2	Decision tree algorithm J48	is a simple C4.5 decision tree for classification. In this type of classifier, a binary tree is constructed to model the classification process. Once the tree is built, it is applied to each tuple in the database and results in classification for that tuple. While building a tree, J48 ignores missing values i.e. the value for that item can be predicted based on what is known about attribute values for the other records. J48 allows classification via either decision trees or rules generated from them. [9]
3	Random forest classifier	is an ensemble learning method for classification that operates by constructing a multitude of decision trees at training time and outputting the class, that is the mode of the classes output by individual trees. In this, the features are randomly selected in each decision split. The correlation between trees is reduced by randomly selecting the features which improves the prediction power and results in higher efficiency. The primary advantage of Random Forest is overcoming the problem of over fitting. [9]
4	Rule Based classifier	is one of the basic and most popular rule based algorithm in which the learned model is represented as a set of IF-THEN rules e.g JRip (RIPPER). Classes are examined in increasing size and an initial set of rules for the class is generated using incremental reduced error. JRip (RIPPER) proceeds by treating all the examples of a particular judgment in the training data as a class, and finding a set of rules that cover all the members of that class. Thereafter it proceeds to the next class and does the same, repeating this until all classes have been covered. [9]

Table two
Characteristics of the study participants

CSID' : Community screening instrument for dementia

	Men (n=286)	Women (n=190)
Age in yrs Mean (SD)	62.03 (5.16)	62.78 (5.41)
Education n(%)		
Illiterate	7 (2.4)	8 (4.2)
Primary	30 (10.5)	30 (15.8)
Secondary	77 (26.9)	62 (32.6)
College	57 (19.9)	41 (21.6)
Diploma	36 (12.6)	23 (12.1)
Graduate	45 (15.7)	8 (4.2)
Post-graduate	34 (11.9)	18 (9.5)
Standard of Living Index Mean (SD)	37.74 (6.95)	36.44 (8.00)
Current occupation n (%)		
Paid full time work	151 (5.28)	17 (8.9)
Paid part time work	19 (6.6)	3 (1.6)
Unemployed	6 (2.1)	3 (1.6)
Housewife/husband	5 (1.7)	137 (72.1)
Retired	105 (36.7)	30 (52.8)
Marital status n (%)		
Never married	7 (2.4)	5 (3.5)
Married/co-habiting	260 (90.9)	115 (60.5)
Widowed	14 (4.9)	68 (35.8)
Divorced	5 (1.7)	2 (1.1)
Global cognitive score Mean(SD)	28.52 (2.50)	28.38 (2.55)
Verbal fluency Mean(SD)	13.93 (3.77)	14.10 (3.83)
Word list recall Mean(SD)	5.0 (1.90)	5.59 (1.99)
Religion n (%)		
Hindu	132 (46.2)	76 (40.0)
Muslim	104 (36.4)	90 (47.4)
Christian and other	50 (17.5)	24 (12.6)
Current smoking n (%)	86 (30.1)	1 (0.5)
Current alcohol n (%)	91 (31.8)	4 (2.1)
	25.13 (4.3)	28.63 (5.4)
BMI Mean (SD)		
Height Mean (SD)	164.76 (6.6)	150.02 (6.23)
Weight Mean (SD)	68.54 (13.6)	64.35 (12.56)
Global cognitive score Mean (SD) (CSID' COGSCORE)	28.52 (2.50)	28.38 (2.55)

	Men (n=286)	Women (n=190)
Verbal fluency Mean(SD)	13.93 (3.77)	14.10 (3.83)
Word list recall Mean(SD)	5.0 (1.90)	5.59 (1.99)

Table three
Assessments for deriving a 10/66 diagnosis of dementia.

Battery of Cognitive tests	<ul style="list-style-type: none"> i) The Community Screening Instrument for Dementia (CSID') COGSCORE incorporating the CERAD animal naming verbal fluency task. (CERAD-Consortium to Establish a register for Alzheimer's Disease) ii) The modified CERAD 10 word list learning task with delayed recall iii) Informant interview, the CSID' RELSCORE, for evidence of cognitive and functional decline
Instruments for diagnosis of dementia	<ul style="list-style-type: none"> i) Battery of cognitive tests (listed above) ii) A structured clinical mental state interview, the Geriatric Mental State, which applies a computer algorithm iii) An extended informant interview, the History and Aetiology Schedule- Dementia Diagnosis and subtype iv) The NEUROEX, a brief fully structured neurological assessment v) Behavioral and Psychological symptoms: assessed by Neuropsychiatric Inventory

Table four
Diagnostic properties of various combinations of cognitive function tests, at different diagnostic cut-offs.

CSID': Community Screening Instrument for Dementia; VF: Verbal Fluency ; WLR: Word List Recall

No	Definition of dementia	Sensitivity	Specificity	Positive Predictive value	Negative predictive value	Kappa
1.	COGSCORE lower than the population mean	0.52 (0.33-0.70)	0.99 (0.98-1.0)	0.74 (0.54-0.93)	0.97 (0.96-.99)	0.59
2.	VF score lower than the population mean	0.78 (0.62-0.93)	0.98 (0.97-0.99)	0.7 (0.54-0.88)	0.99 (0.98-1.0)	0.72
3.	WLR score lower than the population mean	0.93 (0.83-1.02)	0.70 (0.66-.74)	0.16 (0.10-0.22)	0.99 (0.99-1.0)	0.19
4.	COGSCORE /VF/ WLR Score lower than the population mean	1 (1-1)	0.69 (0.64-0.73)	0.16 (0.11-0.22)	1 (1-1)	0.2
5.	COGSCORE &VF/ VF& WLR / Cog score & WLR Score lower than the population mean	1 (1-1)	0.99 (0.98-1.0)	0.84 (0.72-0.97)	1 (1-1)	0.69
6.	COGSCORE lower than 0.5 SD lower than population mean	0.26 (0.09-0.42)	1.0 (0.99-1.0)	0.78 (0.51-1.04)	0.96 (0.94-0.97)	0.37
7.	VF score lower than 0.5 SD lower than population mean	0.56 (0.37-0.74)	1.0 (0.99-1.00)	0.88 (0.73-1.30)	0.97 (0.96-0.99)	0.66
8.	WLR score lower than 0.5 SD lower than population mean	0.74 (0.58-0.91)	0.86 (0.82-0.89)	0.24 (0.15-0.34)	0.98 (0.97-0.99)	0.3
9.	COGSCORE or VF or WLR Score lower than 0.5 SD lower than population mean	0.85 (0.72-0.98)	0.85 (0.81-0.88)	0.26 (0.16-0.34)	0.99 (0.98-1.0)	0.33
10.	COGSCORE &VF or VF& WLR or Cog score & WLR Score lower than 0.5 SD lower than population mean	0.64 (0.43-0.84)	1 (1-1)	1 (1-1)	0.98 (0.97-0.99)	0.76
11.	COGSCORE or VF lower than 0.5 SD lower than population mean & WLR lower than 0.5 SD lower than population mean	0.52 (0.33-0.71)	1 (1-1)	1 (1-1)	0.97 (0.96-0.98)	0.67
12.	COGSCORE lower than 1 SD lower than population mean	0.19 (0.04-0.33)	1.0 (0.99-1.00)	0.71 (0.38-1.05)	0.95 (0.93-0.97)	0.27
13.	VF score lower than 1 SD lower than population mean	0.15 (0.01-0.28)	1.0 (0.98-1.00)	0.67 (0.28-1.04)	0.95 (0.93-0.96)	0.22
14.	WLR score lower than 1 SD lower than population mean	0.37 (0.18-0.55)	0.94 (0.92-0.97)	0.29 (0.14-0.45)	0.96 (0.94-0.98)	0.28
15.	COGSCORE or VF or WLR Score lower than 1 SD lower than population mean	0.44 (0.27-0.63)	0.93 (0.91-0.96)	0.3 (0.15-0.44)	0.96 (0.95-0.98)	0.31
16.	COGSCORE &VF or VF& WLR or Cog score & WLR Score lower than 1 SD lower than population mean	0.18 (0.04-0.33)	1 (1-1)	1 (1-1)	0.95 (0.93-0.97)	0.29
17.	COGSCORE lower than 1.5 SD lower than population mean	0.18 (0.04-0.33)	0.99 (0.99-1.00)	0.71 (0.38-1.05)	0.95 (0.93-0.97)	0.28
18.	VF score lower than 1.5 SD lower than population mean	0.074 (-0.02-0.17)	1 (1-1)	1 (1-1)	0.946 (0.92-0.96)	0.13
19.	WLR score lower than 1.5 SD lower than population mean	0.29 (0.12-0.47)	0.98 (0.97-1.0)	0.57 (0.31-0.83)	0.96 (0.94-0.98)	0.36

No	Definition of dementia	Sensitivity	Specificity	Positive Predictive value	Negative predictive value	Kappa
20.	cog score or VF or WLR Score lower than 1.5 SD lower than population mean	0.37 (0.19-0.55)	0.98 (0.97-0.99)	0.55 (0.33-0.78)	0.96 (0.94-0.98)	0.42
21.	COGSCORE & VF or VF& WLR or Cog score & WLR Score lower than 1.5 SD lower than population mean	0.15 (0.01-0.28)	1 (1-1)	1 (1-1)	0.950 (0.93-0.97)	0.25
22.	COGSCORE lower than 2 SD lower than population mean	0.19 (0.04-0.33)	0.99 (0.99-1.00)	0.83 (0.53-1.13)	0.95 (0.93-0.97)	0.28
23.	VF score lower than 2 SD lower than population mean	0.03 (-0.03-0.11)	1 (1-1)	1 (1-1)	0.94 (0.92-0.96)	0.06
24.	WLR score lower than 2 SD lower than population mean	0.26 (0.04-0.42)	0.99 (0.98-1.00)	0.7 (0.41-0.98)	0.96 (0.94-0.97)	0.35
25.	COGSCORE or VF or WLR Score lower than 2 SD lower than population mean	0.29 (0.12-0.47)	0.99 (0.98-1.0)	0.67 (0.40-0.93)	0.96 (0.94-0.98)	0.38
26.	COGSCORE & VF or VF& WLR or Cog score & WLR Score lower than 2 SD lower than population mean	0.15 (0.01-0.28)	1 (1-1)	1 (1-1)	0.95 (0.93-0.97)	0.24
27.	COGSCORE lower than the population mean	0.52 (0.33-0.71)	0.99 (0.98-1.0)	0.74 (0.54-0.93)	0.97 (0.95-.99)	0.59
28.	VF score lower than the population mean	0.78 (0.62-0.93)	0.98 (0.97-0.99)	0.7 (0.54-0.88)	0.99 (0.98-1.0)	0.72
29.	WLR score lower than 1 SD lower than population mean	0.37 (0.19-0.55)	0.94 (0.92-0.96)	0.29 (0.14-0.45)	0.96 (0.94-0.98)	0.28
30.	COGSCORE or VF Score lower than population mean or WLR score lower than 1 SD lower than population mean	1(1-1)	0.91 (0.89-0.94)	0.41 (0.29-0.53)	1 (1-1)	0.55
31.	COGSCORE & VF or VF& WLR or Cog score & WLR Score lower than specified mean lower than population mean	0.52 (0.33-.71)	1 (1-1)	1 (1-1)	0.97 (0.95-0.99)	0.67
32.	COGSCORE or VF score lower than the population mean and WLR score lower than 1 SD lower than population mean	0.7 (0.53-0.87)	1 (1-1)	1 (1-1)	0.98 (0.97-0.99)	0.81

Table 5
Performance measures after applying Jrip classification for equation 32

colloquium

Classification Jrip	Classification Accuracy (%)		Precision No Units		Recall No Units		Time taken to build the model (in seconds)	
	Training set	Cross Validation	Training set	Cross Validation	Training set	Cross Validation	Training set	Cross Validation
Before Feature Selection	100	100	1	1	1	1	0.06	0.05
After feature selection	100	100	1	1	1	1	0	0