

Short and Long Term Prediction of MCI-to-AD Based on Longitudinal Multimodal, Multivariate Biomarkers

for Alzheimer's Disease NeuroImaging Initiative

Sidra Minhas^{1,2}, Aasia Khanam¹, Farhan Riaz², Shoab A. Khan² and Atif Alvi¹

Abstract— Amongst the 35 million people suffering from dementia, the most prevalent form of dementia is due to Alzheimer's Disease (AD). As of now, no cure to this fatal neurological disorder exists, hence it is important to identify AD at an early stage. The main goal of this work is to identify the Mild Cognitive Impairment (MCI) patients who will convert to AD in coming one or two years or conversely, retain the stable diagnosis of MCI. The proposed framework uses t-test for biomarker ranking. In a wrapper based, 10 fold cross validation system, four types of various dimensioned features are extracted from each of the longitudinal biomarkers. Later three types of binary, supervised classifiers namely, K nearest neighbours, Random Forests and Support Vector Machines are employed to quantize the efficacy of the biomarker-feature subsets. Maximum accuracy of 84.27% and 83.04% is achieved for short-term and long-term conversion prediction respectively using multimodal biomarkers and a high dimensioned feature set.

Keywords— Mild Cognitive Impairment, Alzheimer's Disease, prediction, classification, biomarkers

I. INTRODUCTION

MOST common form of dementia amongst the elderly is due to Alzheimer's disease. No cure or remedies exists as of now for this neurodegenerative disease which has proportionally increased the number of deaths caused by it. In 2012, formal AD staging was performed by the NIA and AA to capture the disease dynamics [1]. It was found that degenerative amyloid AB depositions occur in the pre-symptomatic stage of AD [2]. Next neuronal degeneration occurs which causes mild decline in cognitive abilities [3]. This stage was named as Mild Cognitive Impairment (MCI). The last and final stage is dementia when brain damage becomes extensive and is visible through major cognitive decline, memory problems and inability to perform routine activities [4]. However, it was concluded by Manly et al. [5] that only 10% of the MCI patients convert to AD in the future (MCIc) while the others retain MCI (MCIinc) diagnosis. Hence it is important to identify the patients at risk of progressing to

AD to provide them with better counselling and quality of living. This could be the first step in setting up clinical trials in order to find a remedy for AD.

With advances in medical information gathering, a variety of biomarkers have been suggested for identifying and evaluating AD stages. A very intuitive hypothetical biomarker model with respect to disease stages is presented by Sperling et al. [3] which suggests that cognitive measures and brain morphometric measures obtained through Magnetic Resonance Imaging (MRI) are best suited for detection of MCI-to-AD progression. In the last decade, these two biomarkers in conjunction with several others have been explored by numerous researchers to predict MCI conversion to AD. In [6] [7] [8] [9] [10] several machine learning methods have been applied to multimodal and multivariate cognitive, imaging, plasma and genetic biomarkers. Nonetheless, no consensus has been reached about the best predictors of MCI-to-AD. Recent studies including [11] [12] [13] [10] [14] have experimented with cross-sectional and longitudinal biomarkers yet the accuracy obtained remains limited.

In the present study, baseline (BL) and first annual follow-up (M12) readings of cognitive and MRI biomarkers are employed to predict conversion to AD 1 and 2 years ahead of clinical diagnosis. Four types of features are extracted from each biomarker and classified using three state-of-the-art classifiers. Rest of the paper is organized as follows: Section 2 describes the materials and methods proposed, Section 3 demonstrates the results where Section 4 entails the conclusions.

II. MATERIALS AND METHODS

A. Materials

The data used in this study is derived from the most widely used, open source AD database, the Alzheimer's Disease NeuroImaging Initiative (ADNI). ADNI is a 5 year public-private partnership engaged in gathering AD related biomarkers. We download the longitudinal cognitive and MRI derived measures of the 400 MCI patients recruited in ADNI-1 phase from ADNI website on 30th May, 2015. From amongst these MCI patients, 200 are reported to progress to AD, 100 retained stable MCI diagnosis while the remaining 100 had

¹Department of Computer Sciences, Forman Christian College, Lahore, Pakistan. sidraminhas@fccollege.edu.pk

²Department of Computer Engineering, National University of Sciences & Technology, Islamabad, Pakistan

unstable diagnosis over a period of 3 years. For up-to-date information see www.adni-info.org.

The present study employs 8 cognitive measures that are the scores of the AD Assessment Scale (ADAS), Rey Auditory Verbal Test (AVLT), Clock Drawing Test (CDT), Clock Copying Test (CCT), Immediate Recall Total Score (LIMM), Mini Mental State Examination (MMSE), Trail Making Test A (TRAA) and Trail Making Test B (TRAB). The details of the test procedures and scoring criterion is given in the ADNI General Procedures Manual. In addition to these performance score, MRI derived morphometric measures as provided by the University of California, San Francisco, Memory and Ageing Centre on the ADNI website are used. These morphometric measures include 264 features encapsulating the volumes, surface area and cortical thickness of the brain regions. In this study only those instances are kept in the dataset that have both baseline and first annual follow-up reading available for all biomarkers. Hence the dataset consisted of 73 MCIC and 65 MCInc cases when 1 year ahead prediction was tested and 90 MCIC and 65 MCInc cases when 2 year ahead prediction was made.

B. Methods

The main workflow proposed in this study is shown in Fig. 1. Each module is detailed in the following pages.

1) Biomarker Pre-processing

First of all, the raw ADNI biomarkers are standardized to be used in a machine learning set up. The longitudinal biomarker values are scaled to range between 0 and 1 so as to avoid bias induced by very large and very small values. Later, two-sampled student's t-test is performed to quantify the relative effect of each biomarker's significance in segregating MCInc from MCIC. The p-values of the t-test are used to rank the biomarkers in order of decreasing predictive power. A wrapper based setup is used for biomarker selection in which biomarkers are incrementally considered for classification according to their ranks. Each biomarker subset is evaluated for performance in a cross validation setup.

2) Classification Setup

To assess the performance of our method in a real world setting, we adopted a leave one out cross validation scheme according to which the dataset is divided into disjoint train and

validation set which are run multiple times to provide better estimates of method generalizability. The train and validation set sizes vary according to the size of dataset under consideration.

3) Feature Extraction

In this study, we experimented on four types of feature sets to compare and benchmark our findings. First feature set (FS1) consists of baseline biomarker values only. The second feature set (FS2) consists of both baseline and first annual follow-up readings. However, the evolution in longitudinal biomarker values over time is of interest. Therefore our third feature set (FS3) consisted of the values representing the gradient of annual change in biomarker values. The last three dimensional feature set (FS4) comprised of a combination of time domain biomarker values and the gradient of annual change.

4) Classification

The features obtained are classified into MCIC and MCInc using three types of supervised binary classifiers 1. K Nearest Neighbour, 2) Random Forests, 3) Support Vector Machines (SVM). These classifiers are briefly described below.

a) K Nearest Neighbour (KNN) Classifier:

KNN is an instance based classifier that simply stores all the training data instead of abstracting information from it. Every test instance is evaluated for similarity with the training data. The final class label is assigned as the majority of the labels of 'k' closest neighbors. In this study, the similarity is judged using the Euclidean distance and 'k' is set to 3.

b) Random Forests (RF):

RF is an ensemble classification technique to improve generalizability of a single classifier. In RF classification, a multitude of decision trees are constructed at training time and the class that is the mode of the classes is outputted. In this work, ensembles of J48 tree classifiers are built and the candidate set for number of trees is varied from 5-20.

c) Support Vector Machines (SVM):

SVM is classically employed for numerous classification problems in biomedical domain. SVM algorithm models the features using a kernel function and an optimal

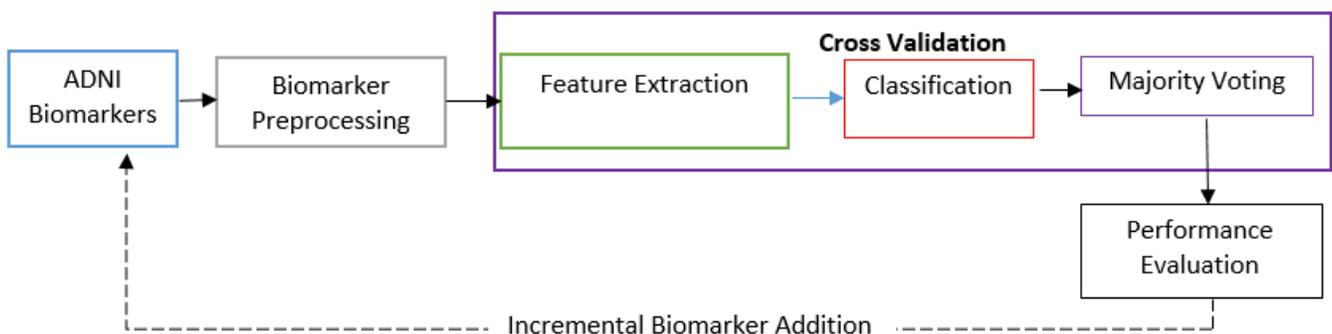


Fig 1. Proposed framework

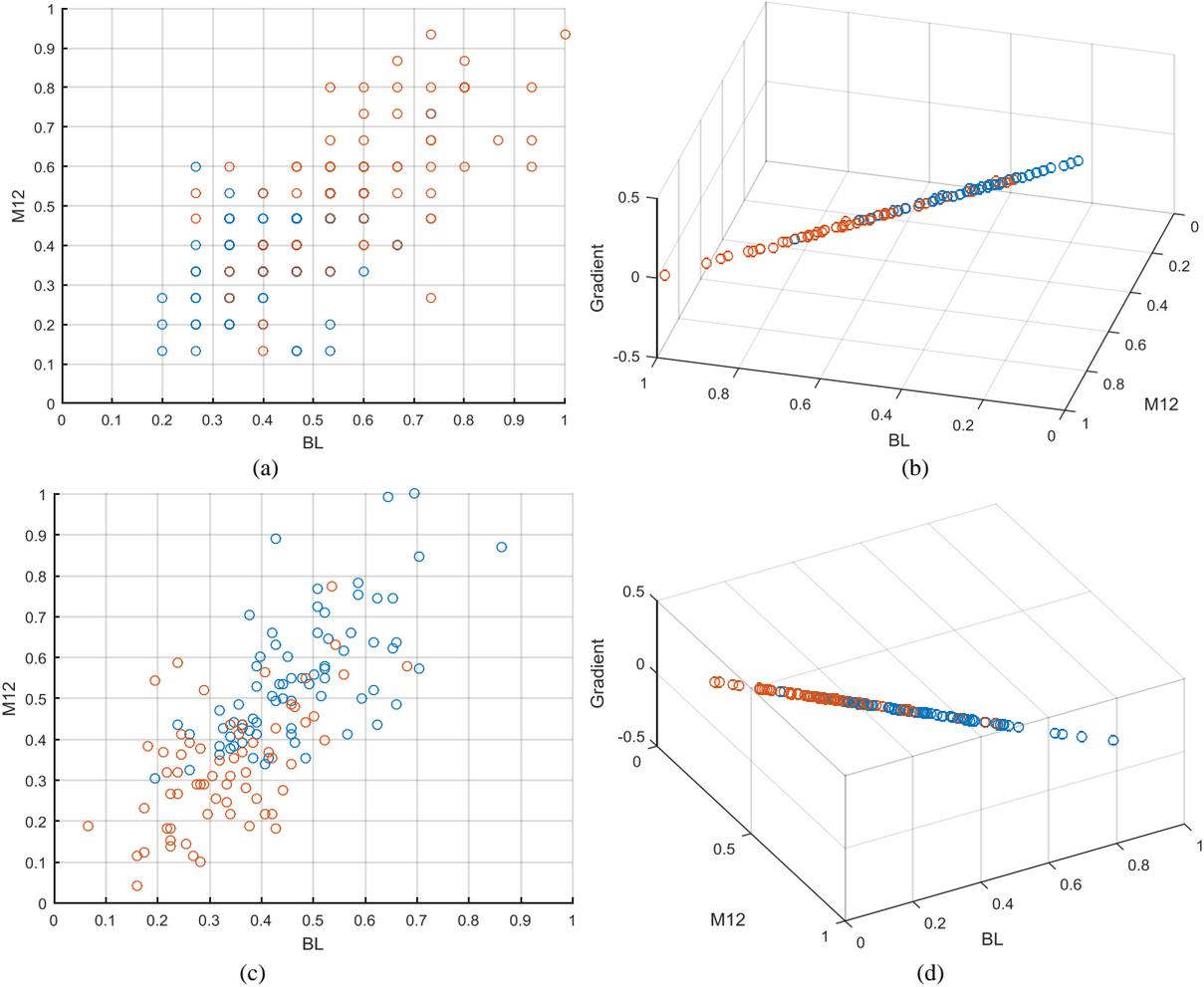


Fig. 2 Pictorial representation of 2D and 3D feature sets of (a) RAVLT FS2, (b) RAVLT FS4, (c) ADAS FS2, (d) ADAS FS4 (Blue MCIInc, Red: MCIc)

hyperplane is fitted such that maximum separation between the two classes is obtained.

5) *Majority Voting:*

For every validation instance, each biomarker is independently used to obtain a class label. Once all biomarkers have generated respective class labels, a majority voting scheme is adopted to assign the final class label to that instance.

6) *Performance Evaluation:*

For every biomarker model and feature set, Area under the ROC curve (AUC) is recorded along with Accuracy, Sensitivity, Specificity and Precision calculated as under:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

Where

- TP (True Positive): Number of correctly identified positive instances
- FP (True Positive): Number of incorrectly identified positive instances
- TN (True Positive): Number of correctly identified negative instances
- FN (True Positive): Number of incorrectly identified negative instances

III. RESULTS

The ranks assigned to the multimodal, multivariate AD biomarkers for 1 year and 2 year ahead conversion prediction are mentioned in Table I. It can be noticed that the cheap and easy to obtain cognitive score are the two top ranked biomarkers for MCI-to-AD progression. The two and three dimensional feature sets derived from the top two ranked biomarker are visually depicted in Fig.2. The effect of increasing the feature set dimensionality can be clearly observed by strong segregation of the two classes.

TABLE I
BIOMARKER RANKS

Rank	1 year ahead	2 year ahead
1	RAVLT	RAVLT
2	ADAS	ADAS
3	Cortical Thickness of Right Entorhinal	Volume of Left Inferior Temporal
4	Volume of right amygdala	Volume of right amygdala
5	Volume of right hippocampus	Surface Area of Left Isthmus Cingulate

Table II summarizes the performance metrics obtained using the various feature sets coupled with supervised classifiers for 1 year ahead conversion prediction. When KNN was employed, FS2, the two dimensional feature set of time point samples presented the best accuracy (77.58%), AUC (77.26%) and precision (76.47%) with a model of 3 top ranked biomarkers. However, when the same task was repeated employing RF of 10 tree classifiers, FS4 with a biomarker model rank 7 displayed same accuracy (77.58%) but a higher AUC (83.72%) and precision (80.53%). This emphasizes that RF classifiers can provide better generalizability over unknown data as compared to KNN. Nonetheless, a noteworthy increase in performance metrics was observed when SVM classifier was employed was a maximum accuracy of 84.78% and AUC of 90.86% with a biomarker model of rank 3.

Similar trends were observed while making two year ahead conversion prediction which are mentioned in Table III. However, in this case, FS4, the three dimensional feature set outperformed all other feature sets under both classifiers. When classification of MCIc vs. MCInc was performed using KNN, maximum accuracy of 75.5% was obtained with a biomarker model of rank 8. But when RF with 15 tree classifiers was used, the accuracy increased by ~3.5% under a biomarker model of rank 7. In this case a significant

employed. A biomarker model of rank 3 resulted in augmented accuracy of 83.04%, AUC of 90.32 and precision of 79.44.

From both Table II and III, it can be observed that under 1 dimensional feature sets (FS1 and FS3), a significant trade-off between sensitivity and specificity was observed which resulted in decreased accuracy and precision measures. It can also be concluded that one year ahead prediction can be made more accurately and precisely based on baseline and first annual follow-up visit as compared to two year ahead prediction.

The effect of increasing biomarker model ranks under SVM classifier is depicted in Fig. 3 for both one and two year ahead conversion predictions. It can be noticed that initially the accuracy measure increases with an increase in model rank, but after a certain complexity is attained, this measure starts decreasing signifying the curse of dimensionality.

IV. CONCLUSIONS

In this paper we presented an in-depth analysis of effective biomarkers for discriminating the MCI patients who are likely to convert to AD in the coming years. We established that cognitive scores are highly informative in quantifying the disease progression and prediction performance is boosted when cognitive scores are used in conjunction with MRI derived measures. However, when number of biomarkers being used for prediction exceed a certain limit, the classification performance decreases. We performed detailed experimentation and found out that longitudinal data are more informative than cross-sectional data when used for slowly progressing diseases such as AD. With our method, a high accuracy of 84.78% and 83.04% was achieved for 1 year and 2 year ahead MCI-to-AD conversion prediction respectively.

TABLE II
PERFORMANCE RESULTS FOR 1 YEAR AHEAD CONVERSION PREDICTION

	KNN				RF				SVM			
	FS1	FS2	FS3	FS4	FS1	FS2	FS3	FS4	FS1	FS2	FS3	FS4
Rank	9	3	11	3	5	2	1	7	4	3	1	3
Accuracy	68.24	77.58	65.22	74.78	74.84	76.15	59.45	77.58	79.73	84.07	52	84.78
AUC	66.87	77.26	63.96	74.58	75.79	80.62	62.34	83.72	86.53	90.65	61.83	90.86
Sensitivity	88.75	90.71	84.28	85.36	79.11	91.61	71.25	79.82	87.68	90.54	100	90.54
Specificity	45	63.81	43.1	63.81	69.29	58.33	46.19	75.71	70.24	76.67	0	78.1
Precision	64.78	76.47	63.08	75.08	74.8	72.46	60.24	80.53	78.53	83.63	52	84.27

increase in performance was noted when SVM was

TABLE III
PERFORMANCE RESULTS FOR 2 YEAR AHEAD CONVERSION PREDICTION

	KNN				RF				SVM			
	FS1	FS2	FS3	FS4	FS1	FS2	FS3	FS4	FS1	FS2	FS3	FS4
Rank	7	3	1	8	1	3	15	7	2	3	7	3
Accuracy	70.29	72.92	64.62	75.5	73.35	77.64	70.99	79.07	74.83	82.42	58.12	83.04
AUC	65.57	69.29	58.72	72.06	77.6	77.42	73.83	79.46	82.92	89.68	64.54	90.32
Sensitivity	92.22	90	95.55	92.2	79.11	89.29	76.61	88.04	95.56	97.78	100	98.89
Specificity	39.29	48.67	21.9	51.9	65.48	65	64.29	69.05	45.95	60.48	0	60.48
Precision	68.1	71.48	63.39	72.89	75.67	77.22	72.65	77.78	71.53	79.24	58.12	79.44

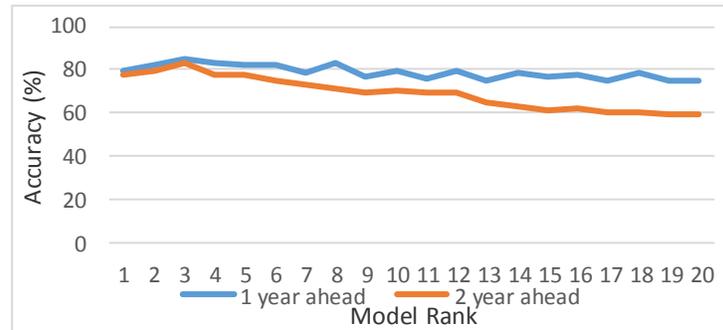


Fig. 3 Variation of classification accuracy wrt to biomarker model ranks

REFERENCES

- [1] H. Braak and E. Braak, "Neuropathological staging of Alzheimer-related changes," *Acta Neuropathologica*, vol. 82, pp. 239-259, 1991.
<http://dx.doi.org/10.1007/BF00308809>
- [2] G. E. Galvin et al., "Practical Guidelines for the Recognition and Diagnosis of Dementia," *Journal of the American Board of Family Medicine*, vol. 25, no. 3, pp. 367-382, 2012.
- [3] R. A. Sperling, P. S. Aisen et al., "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimers & Dementia*, vol. 7, no. 3, pp. 280-292, 2012.
<http://dx.doi.org/10.1016/j.jalz.2011.03.003>
- [4] Markesbery, W. R., "Neuropathologic alterations in MCI: a review," *JAD*, vol. 19, no. 1, pp. 221-228, 2010.
- [5] J. J. Manly, M. X. Tang et al., "Frequency and Course of Mild Cognitive Impairment in a Multiethnic Community," *Ann Neurol*, vol. 63, pp. 494-506, 2008.
<http://dx.doi.org/10.1002/ana.21326>
- [6] Li, F., Tran, L., Han Thung, K. et al., "A Robust Deep Model for Improved Classification of AD/MCI Patients," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1610-1619, 2015.
<http://dx.doi.org/10.1109/JBHI.2015.2429556>
- [7] Davatzikos, C., Bhatta, P., et al., "Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification," *Neurobiology of Ageing*, vol. 32, no. 12, pp. 2322-2327, 2011.
<http://dx.doi.org/10.1016/j.neurobiolaging.2010.05.023>
- [8] Gomar, J. J., Bobes, M. T., et al., "Utility of combinations of biomarkers, cognitive markers and risk factors to predict conversion from MCI to AD in patients in the ADNI," *Arch Gen Psychiatry*, vol. 68, pp. 961-969, 2011.
<http://dx.doi.org/10.1001/archgenpsychiatry.2011.9611>
- [9] Haller, S., Missonnier, P., Herrmann, F. R., Rodriguez, C., Deiber, M.P., et al., "Individual Classification of Mild Cognitive Impairment Subtypes by Support Vector Machine Analysis of White Matter DT," *American Journal of Neuroradiology*, vol. 34, pp. 283-291, 2013.
<http://dx.doi.org/10.3174/ajnr.A3223>
- [10] Lei, B., Ni, D., Wang, T., "Joint Learning of Multiple Longitudinal Prediction Models," *Machine Learning in Medical Imaging*, pp. 330-337, 2015.
- [11] Davatzikos, C., Xu, F., et al., "Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index," *Brain*, vol. 132, pp. 2026-2035, 2009.
<http://dx.doi.org/10.1093/brain/awp091>
- [12] C. Plant, S. J. Teipel, et al., "Automated detection of brain atrophy patterns based on MRI for the prediction of AD," *NeuroImage*, vol. 50, pp. 162-174, 2010.
<http://dx.doi.org/10.1016/j.neuroimage.2009.11.046>
- [13] Moradi, E., Pepe, A., et al., "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects," *NeuroImage*, vol. 104, pp. 398-412, 2015.
<http://dx.doi.org/10.1016/j.neuroimage.2014.10.002>
- [14] Hall, A., Munoz-Ruiz, M, et al, "Generalizability of the Disease State Index Prediction Model for Identifying Patients progressing from MCI to AD," *JAD*, 2014.