Maastricht University

Institute of Data Science, Faculty of Science & Engineering

A clinically translatable machine learning model to predict the conversion from MCI to AD

Nadine Rouleaux[A], Massimiliano Grassi[B], Michel Dumontier[A]

[A] Institute of Data Science (IDS), Maastricht University, Maastricht, The Netherlands
[B] Department of Clinical Neurosciences, Hermanas Hospitalarias - Villa San Benedetto Menni Hospital, FoRiPsi, Italy
Using data from the Alzheimer's Disease Neuroimaging Initiative*

Introduction

Despite the increasing availability in brain health related data, effective methods to predict the conversion from Mild Cognitive Impairment (MCI) to Alzheimer's disease (AD) are still lacking. As currently available and emerging therapies have the greatest impact when provided at the earliest disease stage, the prompt identification of subjects at high risk (e.g. MCI patients) for conversion to full AD is of great importance in the fight against Alzheimer's disease.

Objective

To develop a clinically-translatable algorithm for the identification of Alzheimer's disease conversion in subjects with MCI

•Achieving a high predictive accuracy

•Based only on simple, non-invasive predictors which are either already routinely assessed or effectively introducible in clinical practice

Data Alzheimer's Disease Neuroimaging Initiative

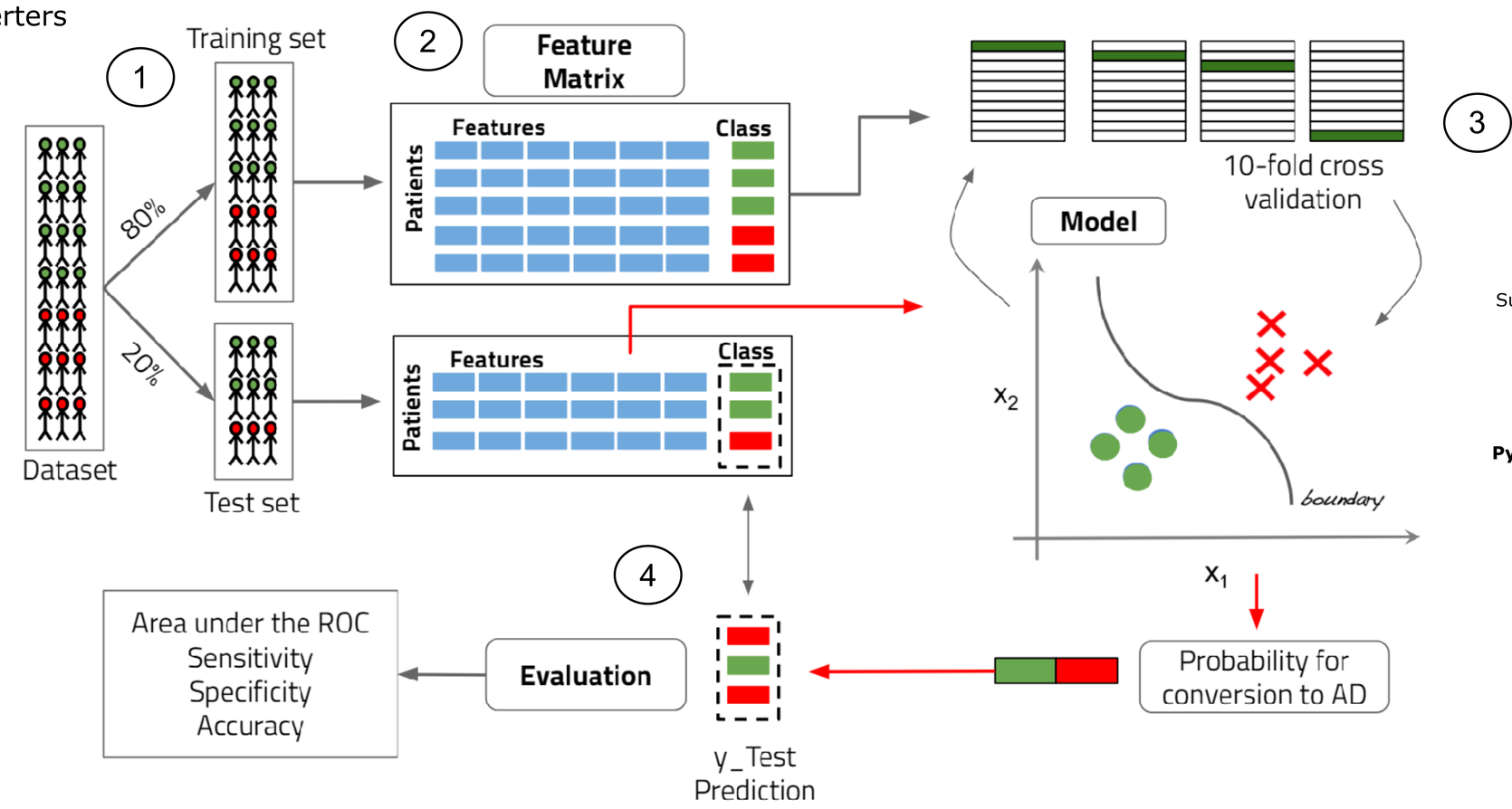
550 MCI patients
197 AD converters
353 AD non-converters

Features

Clinical information

Sociodemographic information

Neuropsychological tests;
Assesment of cognitive functioning



Model Training

10-fold cross-validation, 10 repetitions

100 hyper-parameter Random Search

Models

Logistic regression

Support Vector Machine

K-Nearest Neighbor

Decision Tree

Random Forest

Python 3, sci-kit learn

Preliminary Results

The best model showed a high performance in the cross-validated training set (AUC = 0.876) as well as in the test set (AUC = 0.88).

Conclusion

Our preliminary results show already a high cross-validated and test performance, and the algorithm shows great potential for the prediction of conversion to AD.

We will further optimize the feature sets and hyper-parameter tuning to further improve the performance of our model. Further testing and optimization in independent samples will ensure its translation to clinical practice.

Challenges

Data Quality
•Getting enough data
•Completeness of the features
•The ability to aggregate from different sites (differences in data collection methods, privacy sensitive patient data)

Model Bias
How do you know that the model you learn is appropriate to use on a particular patient? e.g. there could be bias in the population that was learned.

Model Validation
Validation of the model in an independent test set or clinical samples is required. The translation from prototype to application is yet lacking.

** Data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see www.adni-info.org.*

Correspondence to:
Nadine Rouleaux, Maastricht University
Nadine.rouleaux@maastrichtuniversity.nl

Michel Dumontier, Maastricht University
Michel.dumontier@maastrichtuniversity.nl

Institute of Data Science, Maastricht University
www.maastrichtuniversity.nl/IDS

P.O. Box 616
6200 MD Maastricht, The Netherlands