



Applying Deep Learning to Predicting Dementia and Mild Cognitive Impairment

Daniel Stamate, Richard Smith, Ruslan Tsygancov, Rostislav Vorobev, John Langham, Daniel Stahl, David Reeves

► To cite this version:

Daniel Stamate, Richard Smith, Ruslan Tsygancov, Rostislav Vorobev, John Langham, et al.. Applying Deep Learning to Predicting Dementia and Mild Cognitive Impairment. 16th IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI), Jun 2020, Neos Marmaras, Greece. pp.308-319, 10.1007/978-3-030-49186-4_26 . hal-04060674

HAL Id: hal-04060674

<https://inria.hal.science/hal-04060674>

Submitted on 6 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Applying Deep Learning to Predicting Dementia and Mild Cognitive Impairment

Daniel Stamate^{1,2,4}, Richard Smith^{1,2}, Ruslan Tsygancov¹, Rostislav Vorobev¹,
John Langham^{1,2}, Daniel Stahl³, David Reeves⁴

¹ Data Science & Soft Computing Lab, London

² Computing Department, Goldsmiths, University of London

³ Department of Biostatistics and Health Informatics, King's College London

⁴ Division of Population Health, Health Services Research & Primary Care,
University of Manchester

Abstract. Dementia has a large negative impact on the global healthcare and society. Diagnosis is rather challenging as there is no standardised test. The purpose of this paper is to conduct an analysis on ADNI data and determine its effectiveness for building classification models to differentiate the categories Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and Dementia (DEM), based on tuning three Deep Learning models: two Multi-Layer Perceptron (MLP1 and MLP2) models and a Convolutional Bidirectional Long Short-Term Memory (ConvBLSTM) model. The results show that the MLP1 and MLP2 models accurately distinguish the DEM, MCI and CN classes, with accuracies as high as 0.86 (SD 0.01). The ConvBLSTM model was slightly less accurate but was explored in view of comparisons with the MLP models, and for future extensions of this work that will take advantage of time-related information. Although the performance of ConvBLSTM model was negatively impacted by a lack of visit code data, opportunities were identified for improvement, particularly in terms of pre-processing.

Keywords: Dementia prediction, Artificial Neural Networks, Deep Learning, Multi-Layer Perceptron, ConvBLSTM, ReliefF, SMOTE.

1 Introduction

Dementia is a loss of cognitive function. People with this condition have problems with socialization and with their daily lives. The disorder leads to damage or loss of memory, language skills, problem-solving, visual perception, self-management, and the reduced ability to pay attention and to focus. Some people with dementia cannot control their emotions, and their personalities may change. Alzheimer's disease accounts for 60% to 80% of cases of dementia [1].

Worldwide, more than 47.5 million people suffer from dementia currently, and every year there are 7.7 million new cases. The proportion of the general population aged 60 and over with dementia is estimated to be between 5% and 8% [1]. For comparison, dementia is currently more expensive in terms of health care and social assis-

tance than cancer, stroke and chronic heart disease taken together [1]. It is estimated that by 2030, the population with dementia will be approximately 75 million and this condition will cost society US\$ 2 trillion [2]. There is no treatment at the present to cure dementia, however, the life of those patients can be supported and improved. The principal goals for dementia care are: providing information and long-term support to caregivers; optimising cognition, activity, physical health and well-being; detecting and treating behavioural and psychological symptoms; and early diagnosis. In terms of diagnosing dementia, the earlier a person receives a correct diagnosis, the sooner help can be provided. There is however a significant obstacle in diagnosing dementia because there is no standardized test for detecting it [3]. On the other hand, current thinking suggests that 35% of cases of dementia could be prevented if personalized risk could be estimated in advance, helping early informed decisions by clinicians and patients. As such, in addition of efforts to improve dementia diagnosis, there is lately a particular focus of researchers also in developing approaches to predicting risk of dementia [19, 21, 10].

This work proposes a machine learning approach to predicting dementia, in particular based on deep learning. The health care sector is one of the most important areas for machine learning applications. However, it is one of the most complex fields [4] and one of the most challenging, especially in the areas of diagnosis and prediction [5].

According to literature, there are hundreds of possible predictors for dementia, which can generally be categorized based on the following types of models: neuropsychological based models, health-based models, multifactorial models and genetic risk scores [19]. The applicability of these models spreads in multiple directions [16, 20, 21]. The magnetic resonance imaging (MRI), in combination with multiplex neural networks, have been used to segregate healthy brains from progressive mild cognitive impairment (pMCI), based on the structural atrophy of the brain because of Alzheimer's disease [20]. Routine primary care patient records in UK have been and are currently used to develop a risk score for the purposes of estimating how at risk an individual may be of developing dementia, by using conventional statistical methods and modern machine learning algorithms [10, 21]. Positron emission tomography (PET) scans and the regional analysis of the protein amyloid- β , have been used by a Random Forest classifier to identify patients with age-related stable MCI and pMCI [22]. In a recent EMIF-AD study [16], a machine learning methodology based on Extreme Gradient Boosting XGBoost, Random Forest and Deep Learning, has been proposed for Alzheimer's based dementia diagnosis using metabolites in the blood which were proven by the study to be as accurate predictors as the widely accepted but invasive to measure cerebrospinal fluid (CSF) biomarkers.

The present work proposes a new approach to predicting dementia using Deep Learning based on the ADNI (Alzheimer's Disease Neuroimaging Initiative) data [6, 17]. The ADNI dataset we use in this study contains the classes Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and Dementia (DEM). Information gain and ReliefF methods [9, 18] were alternatively used for feature selection. The training, tuning and testing of the predictive models were performed with cross-validation using Deep Neural Network algorithms. The stability of model performances was

studied using Monte Carlo simulations. In particular, this work explores the use of three Deep Learning models: two Deep Feedforward Networks which are Multi-Layer Perceptron models (MLP1 and MLP2), and a Convolutional Bidirectional Long Short-Term Memory (ConvBLSTM) model [7, 8]. Good prediction performances were obtained with MLP1 and MLP2 for dementia: the sensitivities were 0.87 (SD 0.03) and 0.77 (SD 0.01) and the specificities were 0.97 (SD 0.01) and 0.98 (SD 0.01), respectively. The ConvBLSTM model was slightly less accurate in this study but was explored in view of comparisons with the MLP models, and also for future extensions of this work that will take advantage of time-related information in the ADNI data by using the ConvBLSTM’s capabilities exploiting such information.

The remaining of the paper is organised as follows. Section 2 introduces the ADNI data used in this study, and our approach of machine learning methodology based on feature selection, missing data and data imbalance treatments, and deep learning predictive modelling. The prediction results and model performance evaluation are presented in Section 3. Finally a discussion and conclusion section 4 ends the paper.

2 Data and methods

2.1 ADNI data repository

The Alzheimer’s Disease Neuroimaging Initiative (ADNI) data repository was used for this work [17]. The ADNI study was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations as a \$60 million dollar, 5-year, public-private partnership.

There are three main goals of the ADNI study: (1) to detect Alzheimer’s Disease (AD) at the earliest possible stage (pre-dementia) and identify ways to track the disease’s progression with biomarkers; (2) to support advances in AD intervention, prevention and treatment through the application of new diagnostic methods at the earliest possible stages (when intervention may be most effective); and (3) to continually administer ADNI’s innovative data-access policy, which make access to this data possible for scientists worldwide without any embargo.

2.2 The Dataset and the description of variables

The data used in this study was downloaded via the `adnimerge` R package [6], which merges together report forms and biomarkers from the ADNI subprojects ADNIGO, ADNI1, and ADNI2 [17]. The data consists of several different sources: clinical and genetic data, MRI data, PET data and some additional biospecimen. There is also longitudinal information as each participant in the ADNI dataset can have more than one screening visit. In this study we did not work with medical images directly, but used features that had been extracted from the data already.

Our target label consisted of three different classes: Cognitive Normal (CN), Mild Cognitive Impairment (MCI) and Dementia (DEM). The original data comprised 113

variables and 13,272 observations (visits), with multiple observations per participant. The variables extracted from the original dataset were as follows [17]:

- *Baselines Demographics*: age, gender, ethnicity, race, marital status, and education level were included as predictors.
- *Functional Activities Questionnaire (FAQ)* is a test that can be used to assess the dependency on another person that a participant requires to carry out normal daily tasks.
- *Mini-Mental State Exam (MMSE)* is used to estimate the severity and progression of cognitive impairment and to follow the course of cognitive changes in an individual over time.
- *PET measurements (FDG, PIB, AV45)* are participants' brain function measurements.
- *MRI measurements (Hippocampus, intracranial volume (ICV), MidTemp, Fusiform, Ventricles, Entorhinal and WholeBrain)* are structural measurements of participants' brain.
- *APOE4* is an integer measurement representing the appearance of epsilon 4 allele of the APOE gene.
- *ABETA, TAU, PTAU* are cerebrospinal fluid (CSF) biomarker measurements.
- *Rey's Auditory Verbal Learning Test (RAVLT)* are neurophysiological tests evaluating an individual's episodic memory.
- *Everyday cognitive evaluations (Ecog)* are questionnaires that illustrates a participant's ability to carry out everyday tasks.
- *Logical Memory – Delayed Recall Total Number of Story Units Recalled (LDELTOTAL)* is a neuropsychological test that evaluates a person's ability to recall information after a prescribed amount of time.
- *Modified Preclinical Alzheimer Cognitive Composite (mPACC)* are tests that evaluate a person's cognition, episodic memory and timed executive function.
- *ADAS and MOCA* are generalized neuropsychological tests that evaluate a person's cognitive ability (e.g. memory, visuospatial, etc.).

The resulting data contained 1851 participants and 51 input attributes, and one output attribute. 75% of data was used as training data and 25% was used as test data to evaluate the performance of the models.

2.3 Feature selection

To select a set of predictors, two feature selection methods were alternatively employed here. On one hand, we employed the Information Gain method and selected attributes scoring at least 0.01 [18]. On the other hand, we used ReliefF method [9, 18] which we chose to combine with a permutation test based on 500 random permutations of the labels. For instance, features with an observed Relief score corresponding to a distance of at least 1.96 standard deviations from the centre of the normal distribution built with the Relief scores, repeatedly calculated 500 times while labels were permuted randomly, were selected for further processing. This corresponds to the application of the statistical permutation test with significance level $\alpha=0.05$.

Fig. 1 illustrates the example of such a variable mPACCdigit.bl with an observed relief score of 0.38 which is far away from the centre of the distribution, which is an indication that the variable is predictive. The analysis in our study was performed by considering alternative values for the significance level, namely 0.05 and 0.1. However, here we report results obtained with the latter value, which involves a less stringent selection of features, so leads to a larger number of predictors.

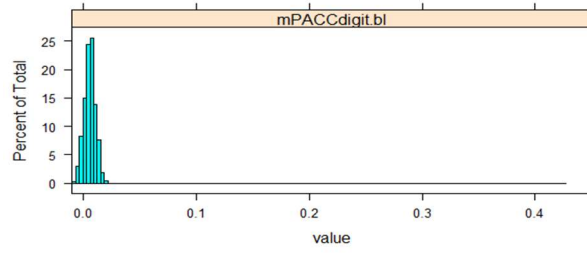


Fig. 1. mPACCdigit.bl variable with an observed relief score of 0.38 which is far away from the centre of the distribution, is considered predictive.

Table 1 illustrates the top 10 ranked features according to the two employed feature selection methods, Information Gain, and ReliefF combined with a permutation test, whose results seem to concord significantly.

Due to lack of space, the predictive modelling results reported in this paper correspond to the second feature selection method only, i.e. ReliefF with permutation test, as this led to better prediction results. Note that mPACCdigit.bl predictor illustrated in Fig. 1 is ranked top by this feature selection method, as illustrated by Table 1. However, the downside of the ReliefF method combined with a permutation test in this study, is that it becomes computationally more expensive.

Table 1: Top 10 features selected with ReliefF and Information gain methods.

Rank	ReliefF	Information gain
1	mPACCdigit.bl	LDELTOTAL.bl
2	mPACCtrailsB.bl	mPACCdigit.bl
3	LDELTOTAL.bl	mPACCtrailsB.bl
4	ADAS13.bl	MMSE.bl
5	FAQ.bl	FAQ.bl
6	MMSE.bl	ADAS13.bl
7	ADAS11.bl	EcogSPTotal.bl
8	EcogSPTotal.bl	ADASQ4.bl
9	EcogSPMem.bl	ADAS11.bl
10	EcogSPPlan.bl	EcogSPMem.bl

2.4 Missing value imputation

For the two MLP models MLP1 and MLP2, 2-dimensional datasets were used, which did not contain longitudinal information. We imputed each visit of a participant from their first visit in order to retain the relationship between different visits and so that the final dataset consisted of only one data instance for each participant.

The ConvBLSTM model used 3-dimensional data, with longitudinal information provided via screening visit data. We replaced the date of each screening with the time difference between visits in order to include information about the sequence of visits as well as the interval between visits. The number of instances of visit codes decreases rapidly after the first visit so the GAIN imputation technique [11] was used to mitigate the lack of later visit data.

2.5 Balancing Classes

Large class imbalance has, most of the times, negative consequences on the performance of the predictive models, as machine learning algorithms tend to focus more on detecting the larger classes. The class distribution for our data is as follows: CN: 617, MCI: 886, Dementia: 348, which suggests that the minority class Dementia could be poorly detected by the predictive models in this case. This problem was mitigated using SMOTE algorithm (Synthetic Minority Over-sampling TEchnique) [12] which reduces class imbalance. SMOTE chooses a data point randomly from the minority class, determines the k nearest neighbours to that point and then uses these neighbours to generate new synthetic data points using interpolation. The synthetic data was added to the minority class to overcome the gap between the majority and minority classes in terms of number of instances. Our analysis used $k=5$ neighbours, as a default value. This technique is to be applied only on the training set as it shouldn't affect the distribution of classes in the test set in order to obtain unbiased estimates for the performance metrics on the test/ unseen data.

2.6 Tuning the Deep Learning models

Parameter values, such as learning rate, decay value and AMSGrad option of ADAM optimiser, for the models were optimised, using Tree of Parzen Estimators (TPE) optimisation algorithm from 'hyperopt' Python package. Models were then evaluated on the test set. Figures 2 and 3 show the topologies of MLP1 and MLP2 optimised models, respectively, which used LeakyReLU activation functions in the hidden layers and a softmax activation function in the output layer. The models were tuned in a 10-fold cross validation using L1 and L2 regularizations, and the dropout regularization with probabilities 0.2 and 0.4, to prevent overfitting.

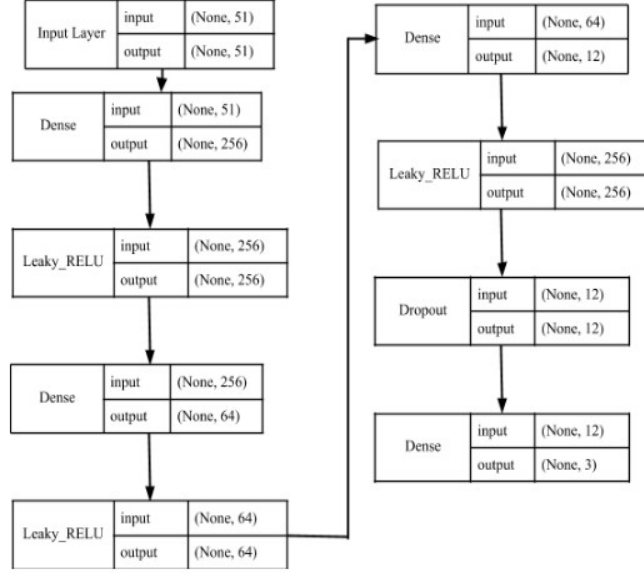


Fig. 2. MLP1: Multilayer perceptron model 1.

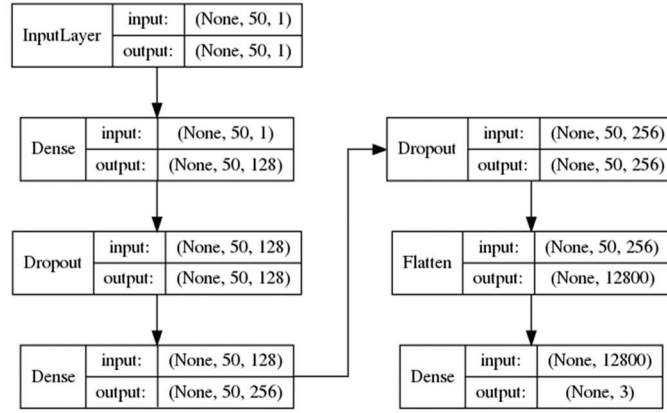


Fig. 3. MLP2: Multilayer perceptron model 2.

The ConvBLSTM model used 3D data, imputed with a generalized GAIN model. The first layers of the model applied convolutional and pooling operations [7], with a kernel that convolutes only operations across different dates for each of the features. Theoretically, this approach could have resulted in a loss of time information but this did not apply in our case because we had a different set of features for different visit codes. The approach also guaranteed that information was preserved for the most recent visit codes with a large proportion of missing values for most persons. Follow-

ing the convolutional operations, we applied a bidirectional LSTM layer [8, 13] because feature distribution helps preserve time information. After that, we applied a flatten operation to the data and sent the result to the final dense layer with softmax activation function, which gave us the probabilities for each target class (DEM, MCI, and CN). Figure 4 shows the topology of the ConvBLSTM model.

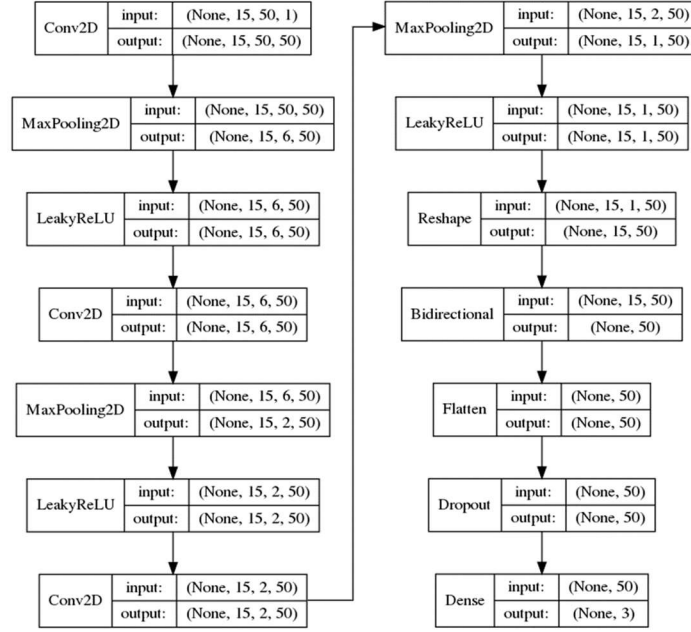


Fig. 4. ConvBLSTM: the Convolutional Bidirectional Long Short-Term Memory model

2.7 Monte Carlo simulation and performance metrics

The stability of well-performing models MLP1, MLP2 and ConvBLSTM was investigated using a Monte Carlo simulation consisting of 100 iterations of the models' tuning and evaluation. We remind that the models' tuning consisted of performing a grid search in a 10-fold cross-validation on the training set (75% of the data) for identifying the best hyperparameters' values. The following performance metrics were evaluated on the test set (25% of the data), and recorded with each Monte Carlo iteration: accuracy, Cohen's kappa statistic, sensitivity (with respect to each class) and specificity (with respect to each class). In addition, for each iteration we recorded the area under the curve (AUC) for each class versus the rest.

2.8 Hardware and software

The MLP1 model was implemented using Python on Tesla k80 GPU, together with a number of packages, including TensorFlow, Keras, hyperopt, scikit-learn, imbalanced-learn, numpy, pandas. The MLP2 model and the ConvBLSTM model were

implemented with Python on Nvidia GTX 1070 Ti GPU under CUDA API. The same Python packages were used as for the MLP1 model, with the exception of Keras, which was used as part of the embedded tf.keras API from TensorFlow.

3 Prediction results

The performance results obtained on the test datasets in the Monte Carlo simulations for our predictive models are presented in this section.

In particular, Fig. 5 below visually compares the boxplots of ROC AUC values for each of the 3 classes MCI, DEM and CN, obtained with 2 conceptually distinct models MLP2 and ConvBLSTM in the Monte Carlo simulation. The boxplots suggest that MLP2 is significantly better than ConvBLSTM if compared with respect to the AUC for MCI class (left side boxplot chart), while the latter model is better than the former model if compared with respect to the AUC for the Dementia class (middle boxplot chart). Moreover, MLP2 is better than ConvBLSTM if compared with respect to the AUC for CN class (right side boxplot chart)

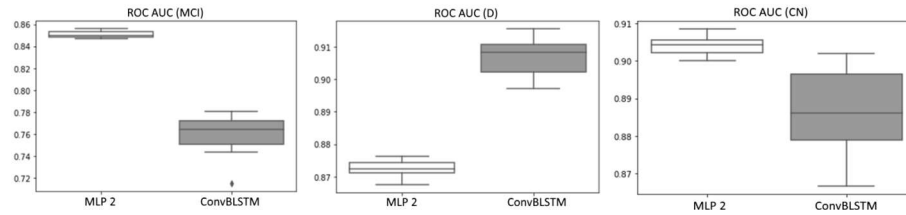


Fig. 5. Monte Carlo: ROC AUC boxplots for Mild Cognitive Impairment (left), Dementia (centre) and Cognitive Normal (right).

Numeric results for the performance metrics in the Monte Carlo simulations for the MLP models MLP1 and MLP2, are presented in Tables 2 and 3, and those for the ConvBLSTM model are presented in Tables 4 and 5. In particular the class-specific performances of the models for dementia, such as AUC, sensitivity and specificity, as well as the non-specific performances accuracy and kappa, are reflected in Tables 2 and 4.

Table 1. Monte Carlo: Accuracy and kappa statistic for all 3 classes (DEM, MCI, CN), and AUC, sensitivity and specificity of Dementia vs all other classes – MLP1 and MLP2

Model	ACC	KAPPA	AUC _{DEM}	SENS _{DEM}	SPEC _{DEM}
MLP 1	0.86 SD 0.01	0.78 SD 0.01	0.92 SD 0.01	0.87 SD 0.03	0.97 SD 0.01
MLP 2	0.85 SD 0.01	0.76 SD 0.01	0.87 SD 0.01	0.77 SD 0.01	0.98 SD 0.01

Table 2. Monte Carlo: Sensitivity and specificity of MCI vs all other classes, and of CN vs all other classes respectively – for MLP1 and MLP2 models.

Model	SENS _{MCI}	SPEC _{MCI}	SENS _{CN}	SPEC _{CN}
MLP 1	0.86 SD 0.03	0.87 SD 0.03	0.94 SD 0.02	0.87 SD 0.04
MLP 2	0.88 SD 0.01	0.83 SD 0.01	0.87 SD 0.01	0.94 SD 0.01

Table 3. Monte Carlo: Accuracy and kappa statistic for all 3 classes (DEM, MCI, CN), and AUC, sensitivity and specificity of Dementia vs all other classes – for the ConvBLSTM model

Model	ACC	KAPPA	AUC _{DEM}	SENS _{DEM}	SPEC _{DEM}
Conv-BLSTM (GAIN)	0.82 SD 0.01	0.72 SD 0.01	0.90 SD 0.01	0.86 SD 0.01	0.94 SD 0.01
Conv-BLSTM (SMOTE)	0.82 SD 0.02	0.71 SD 0.03	0.91 SD 0.01	0.88 SD 0.01	0.93 SD 0.01

Table 4. Monte Carlo: Sensitivity and specificity of MCI vs all other classes, and of CN vs all other classes respectively – for the ConvBLSTM model

Model	SENS _{MCI}	SPEC _{MCI}	SENS _{CN}	SPEC _{CN}
Conv-BLSTM (GAIN)	0.69 SD 0.04	0.86 SD 0.01	0.84 SD 0.03	0.94 SD 0.02
Conv-BLSTM (SMOTE)	0.66 SD 0.04	0.88 SD 0.01	0.87 SD 0.03	0.91 SD 0.03

4 Discussion and conclusion

Deep Learning requires more sophisticated data pre-processing and feature engineering, as well as much more computational time for training a single model in comparison to statistical implementations. The three Deep Learning models explored here were based on two different Deep Feedforward Networks which are Multi-Layer Perceptron (MLP) models, and one Convolutional Bidirectional Long Short-Term Memory (ConvBLSTM) model.

The ReliefF algorithm was the method of choice here used for selecting the most informative features and was a better technique, in this case, in comparison to Information Gain; SMOTE was used to mitigate the large class imbalance in the data.

The ConvBLSTM model was chosen in order to preserve time information in the data. However, the performance of this model was impacted by the lack of visit code data. The GAIN imputation used to mitigate this was suboptimal as the training algorithm made random imputations of the mask layers and projected an initial data distribution on the matrix of data for each person, without considering any inherent differences between various datasets in the ADNI repository. A proposed enhancement for further work is to analyse missing value groups between different visit codes instances and to train the GAIN model for meaningfully different ADNI sub-project data groups, and not for the whole distribution.

The architecture of the neural network based models may be subject of further improvements in forthcoming work. In the case of the ConvBLSTM model, we experimented with a different shape for convolutional operations, so that it would transform participant matrices feature-wise and not time-wise, but this did not lead to a good convergence. An improvement could be to use the convolution operation separately, and apply the transfer learning technique from our already trained model with different convolution operations while freezing the original layers. This would result in a model graph with two different convolution branches and two different bidirectional LSTM layers that would be multiplied before the final dense layer [14].

The magnetic resonance imaging, positron emission tomography and genetic data in the ADNI dataset were only available in numeric and categorical format, and were not as raw images in this study. This meant that the potential power of Convolutional Neural Networks (CNNs) to classify image data was not explored here.

All of the models were able to recognize patterns differentiating the three classes DEM (Dementia), MCI (Minor Cognitive Impairment) and CN (Cognitive Normal), which indicates that each of these machine learning approaches has the capacity to accurately predict dementia and mild cognitive impairment.

Potential directions for further exploration include: (1) improving the modelling of time-based associations, with GAIN imputation to mitigate the sparseness in data for recent visits; (2) exploring other neural network model architectures, such as separating the convolution operation; and (3) applying CNNs to raw image data as these are likely to provide better performance for image classification. With these enhancements it is envisaged that individualized dementia-risk scores will be created and time information will be modelled more effectively.

References

1. World Health Organisation, "Dementia", www.who.int/en/news-room/fact-sheets/detail/dementia.
2. M. Prince, A. Wimo, M. Guerchet, G. Ali, Y. Wu, M. Prina, "The Global Impact of Dementia, An analysis of prevalence, incidence, cost and trends", World Alzheimer Report, 2015.
3. E. Barrett and A. Burns, "Dementia Revealed. What Primary Care Needs to Know", Department of Health UK, 2014.
4. T. Panch, P. Szolovits, and R. Atun, "Artificial intelligence, machine learning and health systems", *Journal of Global Health*, 8(2), 2018.

5. K. Yu, A. Beam, and I. Kohane, "Artificial intelligence in healthcare", *Nature Biomedical Engineering*, 2(10), pp.719-731, 2018.
6. ADNIMERGE, Alzheimer's Disease Neuroimaging Initiative. "ADNIMERGE: Alzheimer's Disease Neuroimaging Initiative". R package version 0.0.1.
7. Shih-Chung B. Lo, Heang-Ping Chan, Jyh-Shyan Lin, Huai Li, Matthew T.Freedman, Seong K. Mun, "Artificial convolution neural network for medical image pattern recognition", 1995.
8. Zhiheng Huang, Wei Xu, Kai Yu, "Bidirectional LSTM-CRF Models for Sequence Tagging", preprint, <https://arxiv.org/abs/1508.01991>, 2015.
9. Newton Spolaôr, Everton Alvares Cherman, Maria Carolina Monard, Huei Diana Lee. "ReliefF for Multi-label Feature Selection", *Brazilian Conf. on Intelligent Systems*, 2013.
10. C. Morgan, D. M. Ashcroft, E. Kontopantelis, D. Stamate, D. Reeves, "Can dementia risk be predicted using routine electronic health records?", *Society for Academic Primary Care Conference (SAPC ASM)*, 2019.
11. Jinsung Yoon, James Jordon, Mihaela van der Schaar. "GAIN: Missing Data Imputation using Generative Adversarial Nets", preprint, <https://arxiv.org/abs/1806.02920>, 2018.
12. N. Qazi, K. Raza, "Effect of Feature Selection, SMOTE and Under Sampling on Class Imbalance Classification", *14th UKSim*, pp. 145-150, 2012.
13. F.A. Gers, J. Schmidhuber, F. Cummins. "Learning to forget: continual prediction with LSTM", *9th International Conference on Artificial Neural Networks*, pp. 850–855, 1999.
14. Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification", *Computers in Biology and Medicine*, Vol. 96, 2018.
15. Karl D. Spuhler, John Gardus III, Yi Gao, Christine DeLorenzo, Ramin Parsey, and Chuan Huang, "Synthesis of Patient-Specific Transmission Data for PET Attenuation Correction for PET/MRI Neuroimaging Using a Convolutional Neural Network", *The Journal of Nuclear Medicine*, pp. 555-560, 2019.
16. D. Stamate, M. Kim, S. Lovestone, C. Legido-Quigley et al., "A metabolite-based machine learning approach to diagnose Alzheimer's-type dementia in blood: Results from the European Medical Information Framework for Alzheimer's Disease biomarker discovery cohort", *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, Vol. 5, 2019, pp 933-938, Elsevier, 2019.
17. Alzheimer's Disease Neuroimaging Initiative <http://adni.loni.usc.edu/>
18. M. Kuhn, K. Johnson, "Applied Predictive Modeling", Springer, 2013
19. L. Robinson, M. Trenell, "Dementia: risk reduction", *Newcastle Inst. of Ageing*, 2016.
20. Nicola Amoroso, Marianna La Rocca1, Stefania Bruno, Tommaso Maggipinto, Alfonso Monaco, Roberto Bellotti, et al., "Brain structural connectivity atrophy in Alzheimer's disease", preprint, arXiv:1709.02369, 2017.
21. K. Walters, S. Hardoon, I. Petersen, S. Iliffe, R.Z. Omar, I. Nazareth, and G. Rait, "Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data", *BMC Medicine*, 2016.
22. D. Brauser, "Machine Learning Tool May Predict Dementia Development Up to 10 Years Later.," *Medscape*, 2016.