Mild Cognitive Impairment Prediction and Diagnostic Procedure Recommendation using Machine Learning

Sajjad Fouladvand, PhD[1], Morteza Noshad, PhD[1], VJ Periyakoil, MD[3], Jonathan H Chen, MD, PhD[1,2]

[1]Stanford Center for Biomedical Informatics Research; [2]Division of Hospital Medicine; [3]School of Medicine; Stanford University, Stanford, CA, USA

Abstract

Without proper prevention and treatment more than 78 million people will be suffering from a type of dementia by 2030. Public health and clinical research initiatives need better means for early identification of patients at risk of dementia, and personalized clinical evaluation steps to diagnose potentially reversible causes. In this study, we leverage real-world electronic health records and an automated machine learning based framework to predict mild cognitive impairment (MCI) as an important risk factor for dementia. Further, our framework includes a recommender system suggesting diagnostic procedures for patients at the risk of MCI as compared to clinical practice. Our experimental results with logistic regression, random forest, XGBoost and long short-term memory models trained and tested on more than 4.2K MCI patients and more than 82K cognitively unimpaired patients show that xgboost model can predict MCI one year before onset of the disease with ROC-AUC of 0.68±0.07 and recommend necessary procedures for MCI patients effectively with average ROC-AUC of 0.78±0.01.

Introduction

Dementia is one of the major causes of mortality and morbidity in older people worldwide and it is estimated that 78 million people will be suffering from some form of dementia by the end of this decade (1), placing a tremendous burden on patients, their families, and health care systems. An important risk factor for dementia is mild cognitive impairment (MCI). Identifying early symptoms of MCI and recommending appropriate diagnostic procedures for patients at the risk of developing MCI is a crucial task to help the aging population with their health needs. Even if there are limited clinical interventions known to effectively alter the course of MCI and dementia, identifying patients at risk would allow for targeted recruitment of such populations into clinical trials to study developing interventions. Learning and disseminating personalized diagnostic evaluation steps is a further essential process to optimize timely diagnosis of MCI cases, including exclusion vs. diagnosis of potentially reversible causes (e.g., endocrine, nutritional, and infectious).

MCI is mainly characterized by minor memory impairment (2) and is formally diagnosed by evaluating individual's cognitive capabilities and clinical examination by a healthcare professional (3). Patients do not routinely screened for possible MCI and as a result are often either under-diagnosed or diagnosis is delayed until late in the illness trajectory. One solution to the lack of formal screening for MCI disease is to identify patients otherwise engaged in the health care system by creating automated tools to analyze patients' medical history and detect those at the MCI risk.

Electronic health records are growing source of information that can be harnessed to identify patients at risk of MCI. Early and accurate diagnosis of such diseases can be addressed using machine learning based tools and analyzing patients electronic health records (EHR) (4,5). To this end, there have been multiple attempts to predict patients with cognitive impairment mostly using standard machine learning models such as support vector machines (SVMs), logistic regression and random forest (6–9) and public databases such as North American Alzheimer's Disease Neuroimaging Initiative (ADNI) (10) as well as the European's AddNeuroMed Study (11). SVM models have been effectively used to predict MCI using gait analysis of patients (12). Other types of healthcare data such as image-based memory test results along with patients' demographics and medical records have been used to produce MCI prediction tools using naïve bayes models (13). More sophisticated deep learning models such as graph convolutional neural networks and recurrent neural networks have also been used to predict MCI onset from patients EHR data as well as imaging and clinical notes data (3,14,15).

However, previous works have been mainly centered on MCI prediction only without providing necessary procedure recommendations for those patients at MCI risk. Further, using machine learning models in MCI prediction has remained largely unexplored mostly due to the lack of real-world data on MCI diagnosis. In this study, we use Stanford EHR data and seek to implement machine learning models for improving dementia diagnosis through the following objectives:

- To determine if machine learning models trained on patients' electronic health records can effectively predict mild cognitive impairment.
- To determine if machine learning models can recommend necessary diagnostic procedures for patients at the risk of developing mild cognitive impairment.

Materials and Methods

The proposed framework in this study includes two main components: 1) MCI onset prediction using machine learning, 2) necessary diagnostic procedure recommendation for patients at the risk of MCI. Figure 1 shows the schema of cohort extraction and model training. Components of this framework are described in detail in the following sub-sections.

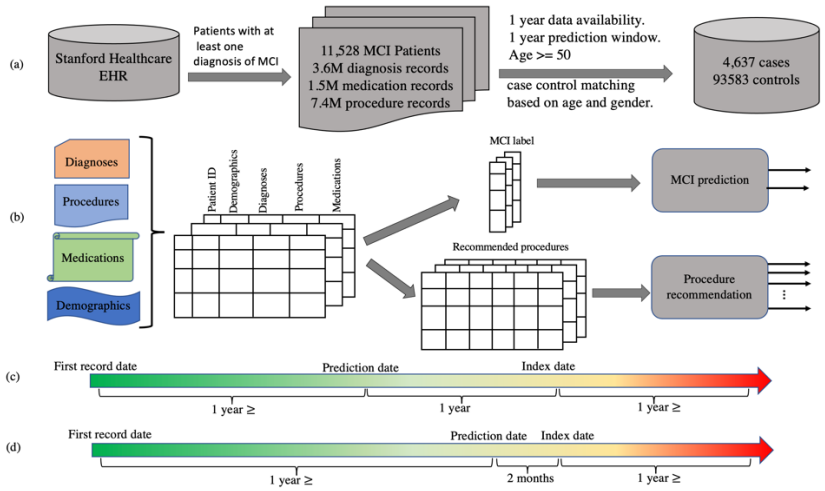

Figure 1. General architecture of our proposed framework for MCI prediction and MCI necessary procedures recommendation. (a) MCI patients were identified, and case and control cohorts were created. (b) Diagnosis, procedure, medication and demographic features were used to create training and testing data for both MCI prediction and procedure recommender system. (c) Timeline for the MCI prediction component: prediction window for MCI prediction is 1 year. Index date for MCI patients is first MCI diagnosis date and for controls is 1 year prior to their last record in the data. (d) Timeline for the diagnostic procedure recommender: prediction window is 2 months.

Data and Cohorts

Our data consist of deidentified EHR records for patients in Stanford Healthcare from 2000 to 2020. Cases include 50 years or older patients with at least one ICD diagnosis of MCI (ICD10s=G31.84, F09 and ICD9s=331.83, 294.9) and at least two years of data availability. Controls are 50 years or older patients with no ICD diagnosis in their records and at least two years of data availability. Data availability is the duration time between patients first record date to the index date where the index date for cases is the first MCI diagnosis date and for the controls is one year before their last record in the data. Controls are matched with cases based on age and gender to create a balanced train set including 5,693 patients (2,840 cases and 2,853 controls) and a test set including 81,078 patients (1,227 cases and 79,851 controls). Our MCI prediction models are trained using balanced training dataset one year prior to the index date for cases and controls; however, our testing experiments include imbalanced scenarios, and the MCI prediction models are tested on the unseen (held out) imbalanced test set. Note, our diagnostic procedure recommender models are trained and tested using Cases only as the recommender system is aimed to provide recommendations for necessary

procedures for patients at MCI risk. Training and testing data for the recommender system include 3,658 and 1,551 patients, respectively.

Data Pre-processing

Dataset includes n patients $P = \{p_1, \dots, p_n\}$. For a patient p_i their diagnosis, medication, procedure and demographic records are in a (t_j^i, C^i) format, where t_j^i is the j^{th} timestamp for p_i and C^i indicate medications, diagnosis or procedures codes for this patient at t_j^i. This longitudinal data format was used to train our long-short term memory model. We converted this longitudinal format to a stationary format for training standard machine learning models including logistic regression, random forest and xgboost. For each patient p_i we computed the frequencies of the features across all patient's timestamps $\{t_j^1, \dots, t_j^i, \dots, t_j^T\}$. We created a stationary dataset $S = \{(X^1, Y^2), \dots, (X^i, Y^i), \dots, (X^n, Y^n)\}$, where X^i is a 1D vector indicating frequencies of all medication, diagnosis and procedure features concatenated with patient's demographic features, age, gender and race, and Y^i is the target variable. For the MCI prediction task Y^i is a binary variable indicating if p_i will develop MCI in a year or not, and for the MCI diagnostic procedure recommendation system Y^i is a multi-hot vector indicating the recommended procedure for p_i. The prediction window for MCI prediction task is one year and for the procedure recommendation task is two months. Table 1 and Table 2 respectively describe statistics of the demographic features and prevalence of the top-5 diagnoses, procedures and medications selected by random forest model. Demographic features in terms of age and gender were similar among cases and controls as we matched based on these variables. The average age of cases and controls were 74.36 years (25th and 75th percentiles = 68, 83) and 76.33 years (25th and 75th percentiles = 70, 84), respectively. The majority were female (54.41% among both cases and controls).

Table 1. Patient demographics among Cases and Controls.

Variable	Case	Control
Age	[a]74.36 (68, 83)	76.33 (70, 84)
Female	2,213 (54.41%)	2,213 (54.41%)
Race		
Asian	[b]458 (11.26%)	595 (14.63%)
Black	225 (5.53%)	150 (3.69%)
Native American	14 (0.34%)	13 (0.32%)
Pacific Islander	32 (0.79%)	32 (0.78%)
White	2,785 (68.48%)	2,446 (60.14%)
Unknown	120 (2.95%)	337 (8.29%)
Other	433 (10.64%)	494 (12.15%)

[a] Numbers are in V(x, y) format, where V is the average and x and y are 25th and 75th percentile, respectively.
[b] Numbers are in N(p%) format, where N is the number of patient and p% shows the percentage in the cohort.

Predictors

Predictors in this study include medications, diagnosis, procedure and demographic features. Note, we grouped medications and diagnoses using pharmaceutical class and clinical classification software (CCS) (16) codes, respectively. We considered the top-100 most frequent medications, diagnoses and procedures codes as well as the demographic features, age, sex and race. The feature set X^i for patient p_i, includes 300 medication, diagnosis and procedure codes plus 3 demographic features. These 303 features were used in our multi label procedure recommender system. We reduced this feature space size further to 30 features using random forests to train machine learning algorithms for our binary MCI prediction system. Random forest and a 5-fold cross validation were used on the training dataset to find top-30 features for the MCI prediction task.

Targets

Target variable for the MCI prediction task is a binary variable indicating if a patient p_i will have an MCI diagnosis code (ICD10s=G31.84, F09 and ICD9s=331.83, 294.9) recorded one year from the time of prediction. Our second

task includes predicting necessary procedures for patients who are diagnosed MCI. The top-100 most frequent procedures among MCI patients were reviewed by our clinical authors to exclude those that reflect low information routine processes (e.g., nursing orders to check vital signs or routine glucose by meter checks) leaving a final target variable set including 46 procedures presented in Table 3. Note, the models were trained using patients' diagnosis, procedure, medication and demographic records up to two months (prediction window in our first task is 12 and in our second task is 2 months) prior to the MCI onset time t_i for each patient p_i. The trained models then predict necessary diagnostic procedures in the form of a multi-hot vector with 46 elements each representing one of the procedures in Table 3.

Table 2. Patient demographics among Cases and Controls.

Variable	Case	Control
Diagnosis Class		
Exposure, encounters, screening or contact with infectious disease	1,537 (37.79%)	565 (13.89%)
Nervous system signs and symptoms	1,332 (32.75%)	379 (9.32%)
Medical examination/evaluation	2,024 (49.77%)	902 (22.18%)
Musculoskeletal pain, not low back pain	1,762 (43.32%)	757 (18.61%)
Disorder of lipid metabolism	2,206 (54.24%)	1,551 (38.14%)
Procedure		
Metabolic panel, comprehensive	2,491 (61.25%)	1,156 (28.42%)
TSH	1,948 (47.90%)	720 (17.70%)
CBC with differential	2,498 (61.42%)	1,279 (31.45%)
Metabolic panel, basic	2,144 (52.72%)	1,055 (25.94%)
Specimen remark	1,544 (37.96%)	674 (16.57%)
Medication Class		
Antihyperlipidemic-hmgcoa reductase inhib(statins)	1,881 (46.25%)	1,060 (26.06%)
Selective serotonin reuptake inhibitor (SSRIs)	793 (19.50%)	246 (6.05%)
Opioid analgesic and non-salicylate analgesics	1,652 (40.62%)	795 (19.55%)
Anticonvulsants	994 (24.44%)	390 (9.59%)
Platelet aggregation inhibitors	1,626 (39.98%)	976 (24.00%)

Numbers are in N (p%) format, where N is the number of patient and p% shows the percentage in the cohort. Non-informative procedures such as external lab results are not shown in this table.

Table 3. List of the diagnosis procedures predicted by the recommender system.

Hemoglobin A1c	Urinalysis, screen for culture
Hepatic function panel	Vitamin b12
Creatinine point of care	CT head
Lipid panel, non-fasting patient	XR chest 1v
Lipid panel, fasting patient	XR chest 2v
Magnesium, serum/plasma	MRI brain wo contrast
Metabolic panel, basic	Blood culture (aerobic & anaerobic bottle)
Metabolic panel, comprehensive	C- reactive protein
Occupational Therapy evaluate	Referral to neurology
Phosphorus, serum/plasma	Eval/mgmt of new patient level 5
Prothrombin time	Eval/mgmt of est patient level 4
Physical Therapy evaluate	Eval/mgmt of est patient level 5
PTT partial thromboplastin time	Autonomic testing- cardiovascular
OT ongoing treatment	PT ongoing treatment
Foley retention catheter	Vitamin d, 25-hydroxyvitamin
Sedimentation rate (esr)	TSH w/ reflex ft4
Referral to physical therapy	Miscellaneous processing
Interagency referral to home health/addendum to certification	CBC w/o diff
T4, free	XR chest 2 views
Referral to neuropsychology	Creatinine, serum/plasma
Troponin i	ECG 12-lead
TSH	Echo - transthoracic echo
Urinalysis, complete	Ferritin

Models

The predictors were used to train logistic regression, random forest, xgboost and long short-term memory (LSTM) models. Logistic regression uses a logistic function to model the outcome probabilities of a single trial experiment(17). Random forest (18) is an ensemble model that operates by constructing a multitude of decision trees at training time and has been used extensively to solve prediction tasks in healthcare data analysis. The goal is to create a predictive model to predict Y^i given the training data set $S = \{(X^1, Y^2), \dots, (X^i, Y^i), \dots, (X^n, Y^n)\}$ of independent random variables distributed as the independent prototype pair $(X^i, Y^i)(15)$. For each tree T_j in a forest including M trees, the predicted value for the input sample X^i is denoted by $m_n(x; \theta_j, D_n)$, where $\theta_1, \dots, \theta_M$ are independent random variables, distributed the same as a generic random variable θ. Similar to random forest, xgboost (19) is an ensample model based on decision trees. XGBoost trains tree ensemble models in an additive manner to greedily and efficiently regularize the ensemble tree objective function. LSTM (20) is a recurrent neural network model where connections between nodes form a directed graph along a temporal sequence and has already been deployed successfully in analyzing temporal data in many biomedical applications(3,21–23). Note, we used LSTM for our second task only, diagnostic procedure recommendation, as LSTM models have shown high capacities in analyzing longitudinal data with multi-label outputs. We used a dynamic LSTM with a fully connected layer and sigmoid activation function on top of the last output of the model to predict necessary procedures for MCI patients (Figure 2).

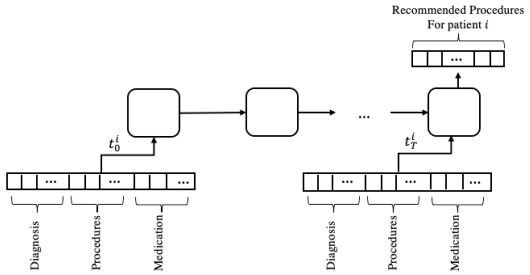

Figure 2. Architecture of the LSTM model for diagnostic procedure recommendation task.

Logistic regression, random forest and xgboost models were trained using the train set including 5,693 (2,840 cases and 2,853 controls) patients' data and a randomized parameter search with a 5-fold cross validation for MCI prediction task. The trained models were tested using a randomly selected held out test set including 81,078 (1,227 cases and 79,851 controls) patients' data. The optimum logistic regression has a sag as it's solver with $C = 100$. The optimum random forest model has 1600 estimators with maximum depth of 110. The optimum values for the number of estimators, maximum depth, learning rate and gamma for the xgboost model used are 1600, 16, 0.1, and 10. The optimized models were assessed using accuracy, precision, recall, F1-score and ROC-AUC. Accuracy is the ratio of correct MCI and CU predictions by the models to the total number of samples in the test set. Precision (positive predictive value) indicators how often the model is correct when predicting a sample as MCI, and recall (sensitivity) measures the performance of the model in retrieving all MCI samples in the test data. F1 is the harmonic mean of precision and recall and ROCAUC shows the model's performance across different decision thresholds. For the procedure recommender task, we used MCI patients' data up to two months prior to their first MCI diagnosis to train and test logistic regression, random forest, xgboost and LSTM models. Similar to the MCI prediction models, the random forest and the xgboost model were optimized using a randomized parameter search with a 5-fold cross validation. However, the LSTM models were trained once using pre-defined default hyper-parameters due to high complexity of fine tuning LSTM models. Note, in this task the models are trained to produce a multi label output indicating necessary procedures (see Table 3) for each patient.

Experimental Results

Table 4 shows the MCI prediction results at one year before the disease onset using logistic regression, random forest and xgboost models. The xgboost model has a slightly higher ROC-AUC (=0.683±0.073) than random forest and logistic regression on the test set. The recall for this model is 0.733, showing that xgboost could correctly detect and predict the majority of MCI patients in the test set at one year before the disease onset. However, precision scores for the xgboost and the other two models are significantly low. This is expected and because of the high control/case ratio in our test set (control/case ratio in the test set is over 65). Precision on a balanced unseen test set including 1,227 cases and 1214 controls for logistic regression, random forest and xgboost were 0.722, 0.762 and 0.780 respectively. Both xgboost and random forest models have ROC-AUC in a range comparable to standard clinical risk stratification and screening models, allowing them to flexible and adjustable based on clinic care settings. They also perform well in detecting majority of patients at the risk of MCI one year before the disease onset.

Table 4. Performance of MCI prediction using machine learning on unseen test sets.

Model	**Precision**	**Recall**	**F1-score**	**AUC**
Logistic Regression	0.023±0.005	0.637±0.130	0.045±0.009	0.676±0.075
Random Forest	0.021±0.003	0.744±0.121	0.040±0.007	0.680±0.078
XGBoost	0.022±0.004	0.733±0.123	0.043±0.007	0.683±0.073

Our second objective in this study is to develop a machine learning model to recommend necessary procedures for patients at the risk of MCI. To this end, we used the MCI cohort (cases) and trained logistic regression, random forest and xgboost models to predict necessary procedures among a list of 46 common diagnostic procedures presented in Table 3. Precision, recall, f1-score and ROC-AUC for these experiments are presented in Table 5. Note, we used random forest, xgboost as well as LSTM models here. XGBoost and random forest both have shown better performance than logistic regression in our first task, MCI prediction and LSTM have showed promising performance in longitudinal data analysis and multi label tasks in healthcare (3,21,23). Further, we used two baseline models: random recommender model and top-10 recommender. Random recommender simply produces random suggestions and top-10 recommender always recommend for the top-10 most frequent diagnostic procedures. We used one-versus-rest micro averaged precision, recall, F1-score and ROC-AUC as the target is a multi-hot vector predicting 46 different diagnostic procedures. Micro averaged assessments were computed by counting the total true positives, false negatives and false positives. XGBoost model had the highest ROC-AUC (0.782±0.017). However, random forest is the most precise model when flagging a patient as high MCI risk (precision=0.594±0.0.091). All machine learning models have higher ROC-AUC than the baseline models, showing the substantial information gain from using personalized prediction models.

Table 5. Machine learning prediction performance in diagnostic procedure recommendation for MCI patients.

Model	Micro averaged precision	Micro averaged recall	Micro averaged f1-score	Micro averaged AUC
Random recommender	0.118±0.016	0.500±0.011	0.191±0.021	0.504±0.010
Top-10 recommender	0.222±0.029	0.409±0.015	0.287±0.026	0.610±0.010
RNN	0.349±0.053	0.204±0.030	0.257±0.037	0.693±0.030
Random forest	0.594±0.0.091	0.062±0.017	0.111±0.029	0.771±0.018
XGBoost	0.565±0.057	0.096±0.017	0.164±0.027	0.782±0.017

Discussion

In this work we analyzed Stanford healthcare EHR data from more than 4.2K MCI and 82K cognitively unimpaired patients over 20 years and created machine learning based models to predict MCI onset as well as creating diagnostic procedure recommendation systems. XGBoost model could predict MCI reasonably effectively (ROC-AUC=0.683±0.073) and produce an effective diagnostic procedure recommendation system for patients at the risk of MCI (ROC-AUC=0.782±0.017). These machine learning models trained using thousands of patients' records could automatically screen patients EHR data and detect those at the risk and suggest initial diagnostic workup steps for those identified. While there are already clinical tools developed to diagnose MCI such as the Montreal cognitive assessment (MoCA) tool (24), these tools are not intended to be *predictive* of future diagnosis, and have typically been produced using small and underrepresented sample size (e.g. only 94 MCI patients data have been used to create MoCA). Further, these tools need to be administered by healthcare professionals with the patient in front of them (or online), which limits their applicability and feasibility. Our automated machine learning based tools may be more feasible to use in large population screening for MCI risk and recommending diagnostic procedures in clinical care than traditional tools.

Limitations in the study include that, although we thoroughly tested the models using randomly selected held-out test sets, more prospective study is needed to test our models' performances in a clinical care environment. Testing both the MCI prediction and diagnostic procedure recommender system in real-time and in a real-world environment can re-assure the generalizability of our models. Further, in this study patients EHR data including their past diagnosis, medication and procedure records were used in a structured format. However, other resources such as patients brain MRI images as well as patients' clinical notes may provide more insights and provide increased power in predicting MCI.

Conclusion

Mild cognitive impairment can be predicted up to 1 year in advance using a combination of structured real-world electronic health record data and machine learning algorithms, and further supported with algorithmically learned diagnostic procedure recommendations for those patients identified.

References

1. Gauthier S, Rosa-Neto P, Morais JA, Webster C. World Alzheimer Report 2021: Journey through the diagnosis of dementia. London: Alzheimer's Disease International; 2021.

2. Robinson L, Tang E, Taylor J-P. Dementia: timely diagnosis and early intervention. BMJ [Internet]. 2015 Jun 16 [cited 2022 Mar 2];350. Available from: https://www.bmj.com/content/350/bmj.h3029

3. Fouladvand S, Mielke MM, Vassilaki M, Sauver JSt, Petersen RC, Sohn S. Deep Learning Prediction of Mild Cognitive Impairment using Electronic Health Records. Proceedings (IEEE Int Conf Bioinformatics Biomed). 2019 Nov;2019:799–806.

4. Goudarzvand S, Sauver JSt, Mielke MM, Takahashi PY, Sohn S. Analyzing Early Signals of Older Adult Cognitive Impairment in Electronic Health Records. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2018. p. 1636–40.

5. Goudarzvand S, St Sauver J, Mielke MM, Takahashi PY, Lee Y, Sohn S. Early temporal characteristics of elderly patient cognitive impairment in electronic health records. BMC Med Inform Decis Mak. 2019 Aug 8;19(Suppl 4):149.

6. Zhang R, Simon G, Yu F. Advancing Alzheimer's research: A review of big data promises. Int J Med Inform. 2017 Oct;106:48–56.

7. Kohannim O, Hua X, Hibar DP, Lee S, Chou Y-Y, Toga AW, et al. Boosting power for clinical trials using classifiers based on multiple biomarkers. Neurobiol Aging. 2010 Aug;31(8):1429–42.

8. Li M, Oishi K, He X, Qin Y, Gao F, Mori S, et al. An Efficient Approach for Differentiating Alzheimer's Disease from Normal Elderly Based on Multicenter MRI Using Gray-Level Invariant Features. PLOS ONE. 2014 Aug 20;9(8):e105563.

9. van Gils M, Koikkalainen J, Mattila J, Herukka S, Lotjonen J, Soininen H. Discovery and use of efficient biomarkers for objective disease state assessment in Alzheimer's disease. Annu Int Conf IEEE Eng Med Biol Soc. 2010;2010:2886–9.

10. Alzheimer's Disease Neuroimaging Initiative [Internet]. [cited 2022 Mar 3]. Available from: http://adni.loni.usc.edu/

11. Lovestone S, Francis P, Kloszewska I, Mecocci P, Simmons A, Soininen H, et al. AddNeuroMed--the European collaboration for the discovery of novel biomarkers for Alzheimer's disease. Ann N Y Acad Sci. 2009 Oct;1180:36–46.

12. Chen P-H, Lien C-W, Wu W-C, Lee L-S, Shaw J-S. Gait-Based Machine Learning for Classifying Patients with Different Types of Mild Cognitive Impairment. J Med Syst. 2020 Apr 23;44(6):107.

13. Bergeron MF, Landset S, Zhou X, Ding T, Khoshgoftaar TM, Zhao F, et al. Utility of MemTrax and Machine Learning Modeling in Classification of Mild Cognitive Impairment. Journal of Alzheimer's Disease. 2020 Jan 1;77(4):1545–58.

14. Zhao X, Zhou F, Ou-Yang L, Wang T, Lei B. Graph Convolutional Network Analysis for Mild Cognitive Impairment Prediction. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). 2019. p. 1598–601.

15. Wang L, Laurentiev J, Yang J, Lo Y-C, Amariglio RE, Blacker D, et al. Development and Validation of a Deep Learning Model for Earlier Detection of Cognitive Decline From Clinical Notes in Electronic Health Records. JAMA Network Open. 2021 Nov 18;4(11):e2135174.

16. HCUP User Support (HCUP-US) [Internet]. [cited 2022 Mar 8]. Available from: https://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp

17. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 2011;12(85):2825–30.

18. Breiman L. Random Forests. Machine Learning. 2001 Oct 1;45(1):5–32.

19. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. New York, NY, USA: Association for Computing Machinery; 2016 [cited 2022 Mar 3]. p. 785–94. (KDD '16). Available from: https://doi.org/10.1145/2939672.2939785

20. Zaremba W, Sutskever I, Vinyals O. Recurrent Neural Network Regularization. arXiv:14092329 [cs] [Internet]. 2015 Feb 19 [cited 2022 Mar 9]; Available from: http://arxiv.org/abs/1409.2329

21. Choi E, Bahadori MT, Schuetz A, Stewart WF, Sun J. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. In: Proceedings of the 1st Machine Learning for Healthcare Conference [Internet]. PMLR; 2016 [cited 2022 Mar 9]. p. 301–18. Available from: https://proceedings.mlr.press/v56/Choi16.html

22. Fouladvand S, Hankosky ER, Bush H, Chen J, Dwoskin LP, Freeman PR, et al. Predicting substance use disorder using long-term attention deficit hyperactivity disorder medication records in Truven. Health Informatics J. 2020 Jun;26(2):787–802.

23. Lipton ZC, Kale DC, Elkan C, Wetzel R. Learning to Diagnose with LSTM Recurrent Neural Networks. arXiv:151103677 [cs] [Internet]. 2017 Mar 21 [cited 2022 Mar 9]; Available from: http://arxiv.org/abs/1511.03677

24. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. Journal of the American Geriatrics Society. 2005;53(4):695–9.