Previsão do comprometimento cognitivo leve: integrando variáveis cognitivas e motoras

Predicting mild cognitive impairment: integrating cognitive and motor features

Predicción del deterioro cognitivo leve: integración de variables cognitivas y motoras

Maria Luiza Wuillaume[1], Jefferson de Moraes Rafael[1], Lucas Martins Lago[1], Jessica Plácido[2], Felipe de Oliveira[3], Pedro Amaral Pereira[1], Manoel Carlos Saisse[4], Claudio Miceli[5], Andréa Deslandes[6] and Andréa Nunes Carvalho[4]

1 Industrial Assessments and Processes Division, National Institute of Technology, Rio de Janeiro, (RJ), Brazil.
2 M.Sc., Institute of Psychiatry, Federal University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil.
3 Prof. M.Sc., Physical Education and Sports Institute, Rio de Janeiro State University, Rio de Janeiro (RJ), Brazil.
4 D.Sc., Industrial Assessments and Processes Division, National Institute of Technology, Rio de Janeiro, (RJ), Brazil.
5 D.Sc, Systems and Computer Engineering Program, Federal University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil
6 D.Sc, Institute of Psychiatry, Federal University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil

Autor correspondente: Ms. Maria Luiza Wuillaume
E-mail: marialuizacw@gmail.com

Resumo

Objetivo: O Comprometimento Cognitivo Leve (CCL) representa uma fase intermediária entre o envelhecimento normal e a demência, exigindo uma detecção precoce para impedir a sua progressão. Este estudo tem como objetivo desenvolver um modelo de classificação de aprendizado de máquina para prever com precisão o prognóstico de indivíduos com CCL, diferenciando-os dos saudáveis.

Método: O método integra variáveis motoras e cognitivas, além de informações autorrelatadas. Foram aplicados os algoritmos SVM, KNN e XGBoost. A melhor previsão foi avaliada pelo método Shapley Value para determinação da importância de cada variável.

Resultados: O SVM apresentou melhor resultado, alcançando 88% de sensibilidade e revelando que as variáveis do domínio motor e dos domínios cognitivo e motor são altamente relevantes para a classificação.

Conclusão: O método desenvolvido, além de ser mais acessível, apresentou alta sensibilidade na classificação do CCL a partir da integração de variáveis cognitivas e motoras.

Descritores: Comprometimento Cognitivo Leve; Aprendizado de Máquina; Prognóstico

Abstract

Objective: Mild Cognitive Impairment (MCI) represents an intermediate stage between normal aging and dementia, requiring early detection to prevent its progression. This study aims to develop a machine learning classification model to accurately predict the prognosis of individuals with MCI, differentiating them from healthy individuals.

Method: The method integrates motor and cognitive variables as well as self-reported information. The SVM, KNN and XGBoost algorithms were applied. The best prediction was evaluated using the Shapley Value method to determine the importance of each variable.

Results: The SVM produced the best results, achieving 88% sensitivity and revealing that the variables of the motor domain and the cognitive and motor domains are highly relevant for classification.

Conclusion: In addition to being more accessible, the method developed also presented high sensitivity in MCI classification based on the integration of cognitive and motor variables.

Keywords: Mild Cognitive Impairment; Machine Learning; Prognosis

Resumen

Objetivo: El Deterioro Cognitivo Leve (DCL) representa un estadio intermedio entre el envejecimiento normal y la demencia, que requiere una detección precoz para evitar su progresión. Este estudio pretende desarrollar un modelo de clasificación de aprendizaje automático para predecir con precisión el pronóstico de los individuos con DCL, diferenciándolos de los sanos.

Método: El método integra variables motoras y cognitivas, así como información autoinformada. Se aplicaron los algoritmos SVM, KNN y XGBoost. La mejor predicción se evaluó mediante el método del valor de Shapley para determinar la importancia de cada variable.

Resultados: El SVM produjo los mejores resultados, alcanzando 88% de sensibilidad y revelando que las variables del dominio motor y de los dominios cognitivo y motor son altamente relevantes para la clasificación.

Conclusión: Además de ser más accesible, el método desarrollado presentó alta sensibilidad en la clasificación de DCL basada en la integración de variables cognitivas y motoras.

Descriptores: Deterioro Cognitivo Leve; Aprendizaje Automático; Pronóstico

Introduction

Mild cognitive impairment (MCI) denotes a transitional phase between typical aging and dementia (1). Early diagnosis of MCI is important to enable early treatment for slowing the progression of dementia (2). Individuals with MCI show a subtle loss, not only in cognitive function, but also in motor performance. For this reason, in addition to cognitive tests, motor features, that can be measured with simple and low-cost tools, have been used as clinical biomarkers to identify MCI (3).

When it comes to diagnosing dementia or MCI using machine learning, studies that developed models with magnetic resonance imaging data produced better results than studies that considered clinical features-based data or voice data (4). Due to the cost and requirement for specialized personnel associated with imaging tests, more recent articles investigated other data sources aiming to achieve similar efficacy, such as self-reported personal information (5) and commonly assessed biomarkers (e.g., glucose, cholesterol, blood pressure) along with demographic and lifestyle information (6).

This research project follows a similar path, by introducing motor features as allies of the cognitive features to build a machine learning classification model that can accurately predict the prognosis of MCI in a more affordable manner. Firstly, a pre-processing stage was conducted to tackle missing features of the dataset, followed by the application of three classification algorithms: Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Extreme Gradient Boosting (XGBoost). In addition, the Shapley Value technique was applied to assess the most important features for the prediction.

The pipeline showcased here is part of a broader research project that aims to develop tools for classifying elderly individuals into one of the three diagnostic categories: MCI, individuals in good health and Dementia. However, the specific emphasis of this work lies on distinguishing between individuals who are healthy from those with MCI. Detecting MCI presents significant challenges due to its subtle nature and the brain's ability to compensate for this impairment, masking its manifestation and leading to its rarity when compared to normal cognitive decline (7).

This paper is organized as follows. Section 2 examines the importance of studying neurodegenerative diseases and how machine learning can address this problem. Section 3 describes the method used to develop the predictive model and evaluate the importance of each dataset feature. Section 4 shows and discusses the results achieved in the research. Section 5 summarizes the pipeline and provides suggestions for future work.

Related Works

Machine learning has been extensively applied as a tool to support early diagnosis of Dementia, Alzheimer's disease, and MCI. Most of those applications rely on image data (8). On the other hand, promising works have been developed utilizing motor and cognitive tests, which are more accessible and less expensive. The authors in Plácido (9) used physical and spatial navigation tests to support screening of Alzheimer and MCI in elderly individuals. Ghoraani (10) proposes an SVM machine learning model to detect MCI and Alzheimer's disease based on dual-task gait assessments applied to a group of 78 individuals.

Health information datasets are prone to missing data, especially in longitudinal studies about the elderly, as their condition may deteriorate during data collection, rendering it impractical to continue with tests and examinations. The work of Okpara (11) assessed this issue and concluded that studies facing it should describe clearly how missing data were addressed. In Minhas et al. (12) a three-step imputation workflow was developed to handle missing data in an aging brain study using machine learning techniques. This workflow outperformed other common methods in terms of prediction and preservation of distribution and correlation among the dataset.

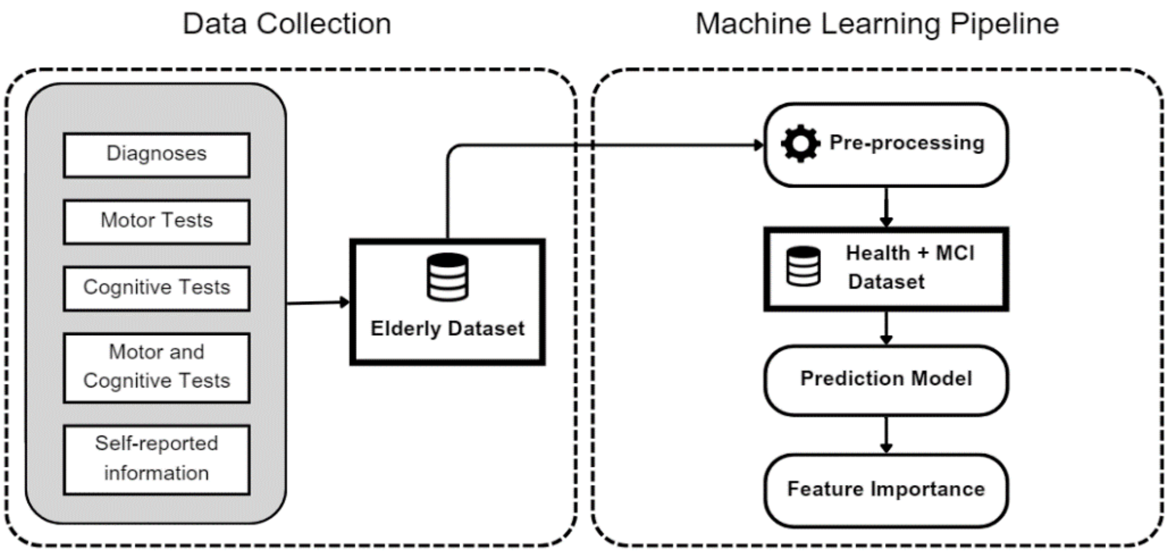
There are no exact formulas to indicate which is the best machine learning algorithm for each specific problem, as the quality of results yielded by each algorithm may vary according to the data characteristics (13). On the other hand, the availability of free machine learning libraries for programming languages, like Python, makes it easier to test different algorithms on the same dataset (14). Wang (15) conducted experiments employing Random Forest (RF), Logistic Regression, Decision Tree, SVM, and XGBoost on a dataset comprising 375 individuals who underwent the mini-mental test for the diagnosis of MCI and dementia. In their dataset, RF yielded the best results in distinguishing between MCI and dementia.

The method presented here aims to classify elderly individuals as either healthy or with MCI based on a broader range of cognitive and motor tests than those found in the reviewed literature. To the best of our knowledge, no prior studies have proposed machine learning models that integrate motor performance and cognitive performance at this level.

Methods

This section describes the method used to develop the classification model that distinguishes healthy patients from those with MCI. Figure 1 illustrates its outline, which can be summarized into 4 main stages: data collection, pre-processing, predictive model development and identification of the most important features.

Figure 1 – Method Outline



Data Collection

For this study, data was gathered from different domains to create a dataset consisting of 272 elderly individuals. Among them, 64% were identified as healthy, 14% were diagnosed with MCI, and 22% were diagnosed as dementia patients. The features that compose the dataset are listed below.

- Five represent the outcomes of tests focused on motor performance, which are:
  1. Berg Balance Test (BBS) (16)
  2. Handgrip Strength (17)
  3. Sit-to-stand Test (STS) (18)
  4. 2-minute Aerobic Test (STEP) (18)
  5. Gait Speed on Single-tasks (GSST) (19)
- Six correspond to the result of tests focused on cognitive performance, which are:
  1. Mini-Mental State Examination (MMSE) (20)

2. Clock Draw Test (CDT) (21)
3. Trail Making Test Part A (TMTA) (21)
4. Digit Span Forward (DSF) (22)
5. Digit Span Backward (DSB) (22)
6. Verbal Fluency (VF) (21)

- Four correspond to the result of tests that assess both cognitive and motor performance, which are:

1. Gait Speed on Dual-tasks (GSDT) (19)
2. Immediate Time on Floor Maze Test (FMT-IT) (18)
3. Delay Time on Floor Maze Test (FMT-DT) (18)
4. Delta Gait Speed (delta GS - calculated subtracting GSST from GSDT)

- Five encompass self-reported information, which are:

1. Schooling
2. Number of Comorbidities
3. Instrumental Activity of Daily Living (IADL) (20)
4. Age
5. Gender

These characteristics arose following a thorough examination of existing literature, aiming to identify tests that were both cost-effective and endorsed by the scientific community.

The Research Ethics Committee approved this cross-sectional study at the Psychiatry Institute of the Federal University of Rio de Janeiro (IPUB-UFRJ) under the CAAE registry: 42349815.0.0000.5263.

Pre-processing

First, the dataset was divided into training records, which correspond to 80% of the data, and test records, which represent 20%.

As there were records with missing data in the training set, a data imputation technique was applied to minimize the loss of records, due to the small sample size. The algorithm chosen for this task was the Iterative Imputer, known as MICE (Multiple Imputation by Chained Equation) (16). It starts by initializing the missing values using an initial estimation method like mean or median. Then, in each iteration, it selects one feature with missing values and treats it as the target feature, using the other features as predictors in a regression model. MICE was selected as a result of prior exploratory analyses and discussions with health researchers, which indicated a potential relationship among the features that regression models could explore to impute missing data.

In the imputation process, all the records from the training set were used to load more information into the regressions. However, for the next steps, the training and test data were filtered to include only healthy patients and patients with MCI. This is due to the fact that this work focuses on the distinction between MCI and healthy patients. The model produced here will be reused in a forthcoming study, which aims to perform a prognosis of the three classes, including dementia.

The dataset was fed into the classification models under two conditions: with and without feature normalization, in order to assess the impact of this process on their performance, particularly for SVM, which can be sensitive to data scaling. When applied, normalization was conducted before the data imputation and was achieved through the utilization of a min-max scaler (23).

Prediction Model

During the classification phase, three different algorithms were evaluated: KNN (23), XGBoost (23) and SVM (23).

The algorithms hyperparameter's were optimized using 5-fold cross-validation. This technique helps to ensure that the chosen hyperparameters generalize well to unseen data by repeatedly partitioning the dataset into training and validation subsets, reducing the risk of overfitting and enabling a more reliable evaluation of the model's effectiveness. For SVM and KNN, a GridSearch strategy was applied to systematically explore various combinations of hyperparameters. In the case of XGBoost algorithm, which encompasses a wide range of hyperparameters, an even more advanced approach was used: Bayesian search (24). This technique considers the probability of different hyperparameters configurations, allowing for a more intelligent exploration of the search space.

Given the unbalanced nature of the data, accuracy was discarded as an optimization metric. Instead, metrics such as f1, f-beta-2 and recall were preferred. Since the goal was to minimize misclassifications of MCI individuals as Healthy, optimizing using f-beta-2 and recall led to improved outcomes by effectively reducing false negatives.

Feature Importance

Gaining insight into which features are most pertinent to the classification process constitutes a crucial aspect of this pipeline. Firstly, it involves determining whether motor aspects genuinely influence MCI prediction. Secondly, it entails discerning which tests are indispensable for prediction and which ones can be omitted, thereby saving time and enhancing accessibility to the process.

The solution selected to produce the feature importance rating was Shapley Value (25). In the context of machine learning, Shapley Value is a technique for model explainability. It assigns a value to each feature in a predictive model, indicating the specific contribution of that feature to the prediction of an outcome. With the contribution values produced by Shapley, it was possible to construct an importance ratio of the different feature domains.

Results and Discussions

The experiments were implemented using the Python programming language supported by the libraries: Scikit-learn, Shap and Xgboost.

Table I presents the best results obtained by each of the tested algorithms, with and without feature normalization. It reveals that, although XGBoost and KNN have shown interesting results with normalized features, SVM applied to non-normalized features stood out for producing the highest scores.

Table I also shows that the SVM results are better when the features are not normalized. This means that some of the most informative features were also those with the largest scale. The Shapley Value evaluation confirmed that these features did, in fact, make a greater contribution to the prediction.

Tabela 1 – Performance metrics according to XGBoost, KNN and SVM models, considering normalized and non-normalized features.

	Non-normalized features			
Classifier	**accuracy**	**f1**	**precision**	**recall**
XGBoost	0.78	0.27	0.28	0.25
KNN	0.71	0.12	0.11	0.12
SVM	0.86	0.67	0.54	0.87
	Normalized features			
Classifier	**accuracy**	**f1**	**precision**	**recall**
XGBoost	0.82	0.40	0.43	0.37
KNN	0.80	0.37	0.37	0.37
SVM	0.84	0.43	0.50	0.37

Figure 2 displays a confusion matrix, which demonstrates that the best model correctly identified 88% of people with MCI (i.e. sensitivity). It also correctly identified 86% of healthy individuals (i.e. specificity). In order to improve the next study, health researchers should analyze patients with MCI who have been misclassified, to understand whether they have the characteristics of healthy individuals for the features considered. This would indicate the need to include additional features to enhance the predictive model.

Figure 3 illustrates the contribution of each domain to the prediction made by the SVM model with the non-normalized features. To produce it, the average absolute value of the features contribution in each of the predicted outcomes was obtained using the Shapley Value evaluation. Then the average values from the features that belonged to the same domain were added together to represent the domain's contribution. Therefore, the percentages shown in the pie chart correspond to the value of the domain's contribution in proportion to the total contribution.

The interpretation of Figure 3 suggests that the features that best supported the prediction belong to the cognitive and motor domain. As expected, the variables with a cognitive focus take a sizable slice of the contribution, but the variables with a motor focus are not far behind either, occupying a significant portion of the graph.

In these outcomes, motor performance proved to be a great ally of cognitive performance in the development of a classifier that reached relatively high sensitivity.

Figure 2 – Confusion matrix - SVM with non-normalized features

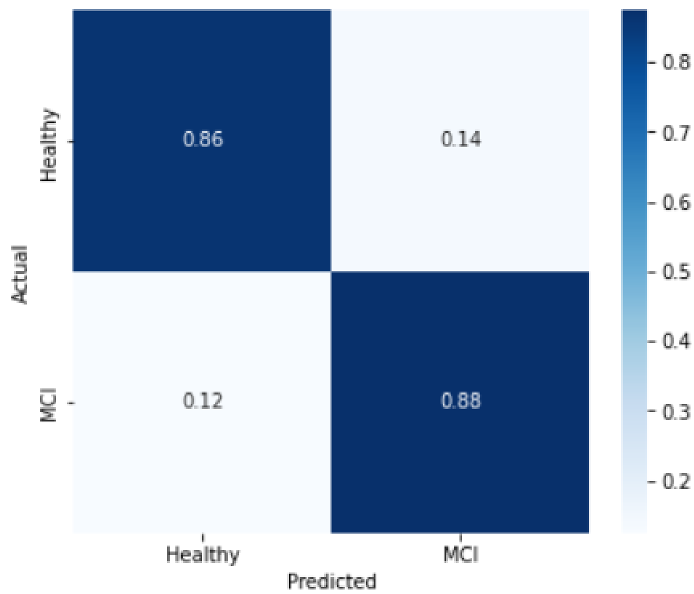
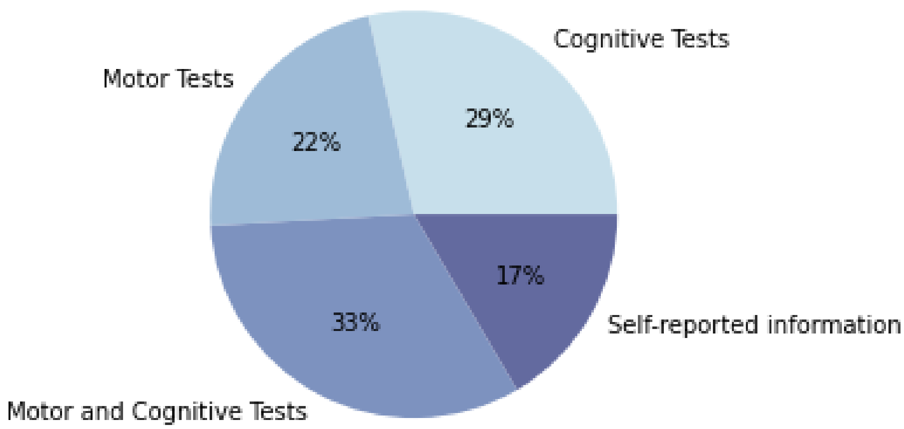


Figure 3 – Relative Contribution of Domains to Prediction



Conclusion

The objective of this study was to develop a machine learning classification model that accurately predicts the prognosis of individuals with MCI, differentiating them from healthy individuals. The constructed machine learning pipeline investigates the fusion of features from different domains to create an efficient and cost-effective method. For this purpose, data was gathered from 272 elderly individuals, including outcomes from tests assessing motor, cognitive, and combined cognitive and motor performance. Additionally, self-reported information was integrated into the data collection process. During the pre-processing stage, the MICE data imputation technique was utilized to address missing features in the dataset. Subsequently, SVM, KNN, and XGBoost algorithms were employed to develop the classification models. Moreover, the best prediction was assessed using the Shapley Value method to determine feature importance.

The SVM model outperformed the others, achieving an 88% sensitivity. It highlighted the importance of features from both the motor domain and the combined cognitive and motor domain in the classification process. The key predictors for precise forecasts predominantly come from the cognitive and motor domains. As expected, variables related to cognition make a substantial impact, while those emphasizing motor skills also contribute significantly, highlighting their relevance in the analysis.

As future works, the research team is collecting new records to increase the dataset. The machine learning models will then be retested with larger patient datasets. Thus, the results will be refined using new validation data to select the best models for the implementation phase.

Acknowledgements

The authors are grateful to National Institute of Technology (INT), to National Council for Scientific and Technological Development (CNPq - 303474/2019-0) and to Rio de Janeiro State Foundation to Support Research (FAPERJ - E26/0.10.001920/2019 - E26/202.523/2019). They are also grateful to Karla Figueiredo, D.Sc. (Institute of Mathematics and Statistics of the State University of Rio de Janeiro - UERJ) for her guidance during the project and to José Vinicius Alves Ferreira, M.Sc. (Institute of Psychiatry, Federal University of Rio de Janeiro) for his help with data collection.

References

1. Petersen RC. Mild cognitive impairment. CONTINUUM: Lifelong Learning in Neurology. 2016;22(2 Dementia):404.

2. Sabbagh MN, Boada M, Borson S, Chilukuri M, Doraiswamy PM, Dubois B, et al. Rationale for Early Diagnosis of Mild Cognitive Impairment (MCI) supported by Emerging Digital Technologies [Re- view]. Journal of Prevention of Alzheimer's Disease. 2020;7(3):158 – 164.

3. de Oliveira Silva F, Ferreira JV, Plácido J, Chagas D, Praxedes J, Guimarães C, et al. Stages of mild cognitive impairment and Alzheimer's disease can be differentiated by declines in timed up and go test: A systematic review and meta-analysis [Review]. Archives of Gerontology and Geriatrics. 2019;85.

4. Javeed A, Dallora AL, Berglund JS, Ali A, Ali L, Anderberg P. Machine Learning for Dementia Prediction: A Systematic Review and Future Research Directions [Article]. Journal of Medical Systems. 2023;47(1).

5. Gómez-Ramírez J, Ávila Villanueva M, Fernández-Blázquez MA. Selecting the most important self-assessed features for predicting conversion to mild cognitive impairment with random forest and permutation- based methods [Article]. Scientific Reports. 2020;10(1).

6. Twait EL, Andaur Navarro CL, Gudnason V, Hu YH, Launer LJ, Geer- lings MI. Dementia prediction in the general population using clinically accessible variables: a proof-of-concept

study using machine learning. The AGES-Reykjavik study [Article]. BMC Medical Informatics and Decision Making. 2023;23(1).

7. Aljumaili M, McLeod R, Friesen M. Serious Games and ML for Detecting MCI. In: 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP); 2019. p. 1-5.

8. Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. Alzheimer's research & therapy. 2021;13:1- 29.

9. Plácido J, Ferreira JV, Araújo J, Silva FdO, Ferreira RB, Guimarães C, et al. Beyond the Mini-Mental State Examination: The use of physical and spatial navigation tests to help to screen for mild cognitive impairment and Alzheimer's disease. Journal of Alzheimer's Disease. 2021;81(3):1243-52.

10. Ghoraani B, Boettcher LN, Hssayeni MD, Rosenfeld A, Tolea MI, Galvin JE. Detection of mild cognitive impairment and Alzheimer's dis- ease using dual-task gait assessments and machine learning. Biomedical signal processing and control. 2021;64:102249.

11. Okpara C, Edokwe C, Ioannidis G, Papaioannou A, Adachi JD, Thabane L. The reporting and handling of missing data in longitudinal studies of older adults is suboptimal: a methodological survey of geriatric journals. BMC medical research methodology. 2022;22(1):122.

12. Minhas S, Khanum A, Riaz F, Khan SA, Alvi A. Predicting Progression From Mild Cognitive Impairment to Alzheimer's Disease Using Autoregressive Modelling of Longitudinal and Multimodal Biomarkers. IEEE Journal of Biomedical and Health Informatics. 2018;22(3):818-25.

13. Sarker IH. Machine learning: Algorithms, real-world applications and research directions. SN computer science. 2021;2(3):160.

14. Bloice MD, Holzinger A. A tutorial on machine learning and data science tools with python. Machine Learning for Health Informatics: State-of-the-Art and Future Challenges. 2016:435-80.

15. Wang J, Wang Z, Liu N, Liu C, Mao C, Dong L, et al. Random forest model in the diagnosis of dementia patients with normal mini- mental state examination scores. Journal of personalized medicine. 2022;12(1):37.

16. Tatiyama Miyamoto S, Júnior I, Berg K, Ramos L, Natour J. Brazilian version of the Berg Balance Scale. Brazilian journal of medical and biological research = Revista brasileira de pesquisas médicas e biológicas / Sociedade Brasileira de Biofísica [et al]. 2004 09;37:1411-21.

17. Teixeira I, de Oliveira Silva F, Ferreira J, Plácido J, Marinho V, Deslandes A. Utility of handgrip strength cut-offs for identification of weakness and disability in community-dwelling older people with mild cognitive impairment and Alzheimer's disease. Jornal Brasileiro de Psiquiatria. 2019 12;68:208-14.

18. Plácido J, Ferreira J, Araújo J, de Oliveira Silva F, Ferreira R, Guimarães C, et al. Beyond the Mini-Mental State Examination: The Use of Physical and Spatial Navigation Tests to Help to Screen for Mild Cognitive Impairment and Alzheimer's Disease. Journal of Alzheimer's Disease. 2021 04;81:1-10.

19. Whitson HE, Potter GG, Feld JA, Plassman BL, Reynolds K, Sloane R, et al. Dual-task gait and Alzheimer's disease genetic risk in cognitively Normal adults: a pilot study. Journal of Alzheimer's Disease. 2018;64(4):1137-48.

20. Han G, Maruta M, Ikeda Y, Ishikawa T, Tanaka H, Koyama A, et al. Relationship between Performance on the Mini-Mental State Examination Sub-Items and Activities of Daily Living in Patients with Alzheimer's Disease. Journal of Clinical Medicine. 2020 05;9:1537.

21. Carvalho G, Caramelli P. Normative data for middle-aged Brazilians in Verbal Fluency (animals and FAS), Trail Making Test (TMT) and Clock Drawing Test (CDT). Dementia Neuropsychologia. 2020 03;14:14-23.

22. Lichtenberger EO, Kaufman AS. Essentials of WAIS-IV assessment. vol. 50. John Wiley & Sons; 2009.

23. Ha J, Kambe M, Pe J. Data Mining: Concepts and Techniques. Elsevier; 2011.

24. Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. vol. 4; 2012. p. 2951 – 2959.

25. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. vol. 2017-December; 2017. p. 4766 – 4775.