Dementia Subtype Prediction at Early Stages using Machine Learning and Structured Data

Hui Wei DAVIDWEI@NYU.EDU
NYU Langone Medical Center

Narges Razavian NARGES.RAZAVIAN@NYULANGONE.ORG
NYU Langone Medical Center

1. Background

Dementia is heterogeneous and currently irreversible condition, that can manifest as Alzheimer's disease (AD), Lewy body dementia (LBD) and other subtypes. Treatments available today can only slow the progression, so early detection is necessary for impaired patients. The Lewy body variant of Alzheimer's disease (Mix AD+LBD) is a less prevalent but more severe concomitant combining AD and LBD. Due to overlapping symptoms among dementia subtypes, clinical diagnosis of the subtype particularly at early stages is difficult, and true nature of dementia can only be confirmed via brain autopsy, post-mortem. In this work, we examined machine learning models to differentiate pure AD, pure LBD, Mix AD+LBD and other subtypes at the first visit of mild cognitive impairment. Unlike prior works [4, 8, 6, 3, 5, 2, 1, 7] that have used clinicians' diagnosis as the basis for analysis, we use autopsy-confirmed neuropathological results as ground-truth labels. To our best knowledge, this is the first work using machine learning models and structured data to differentiate pure AD, pure LBD, Mix AD+LBD and other subtypes based on data captured at patient's first visit with mild cognitive impairment.

2. Methods

This study is based on publicly available data from National Alzheimer's Coordinating Center[1]. We used non-imaging structured data: demographics, health history, neuropsychological tests, CSF values, APOE genotypes as our inputs and autopsy-based neuropathological findings as the label. Autopsy-confirmed labels are based on ABC score and Lewy body pathology, while clinician diagnoses are based on presumptive etiologic diagnosis of AD and LBD. We included patients who have at least one visit with documented mild dementia (global CDR score=0.5 or 1) and autopsy based neuropathological results. Variables from first visits of mild cognitive impairment are used as inputs. We split patients into non-overlapping training, validation and test set (1127:380:380). Multiclass Logistic Regression (LR) and Multilayer Perceptron (MLP) are trained and validated. L1 regularization are used to avoid over-fitting. To select the best model to compare with clinicians, (1) we tuned hyperparameters of each model based on the best Macro and Micro F1 score, (2) among all the tuned models, we removed overfitting ones, and selected one model with the best overall Macro F1 score and one with the best overall Micro F1 score. (3) the overall best model which has lowest validation loss is selected from these two models. We report final results on the held-out test set.

3. Results

Our best model: LR with L1 coefficient of 0.01 has the best Micro F1 score and lowest loss on the validation set. We compare it with clinicians using bootstrapped (95% confidence interval) held-out test set. For F1 score, sensitivity, specificity and confusion matrix of the best model and clinicians, please see Table 1, 2, 3 and Figure 1. The best model outperforms clinicians on Mix AD+LBD, while clinicians outperform the best model on pure LBD and other subtypes. Their performance on pure AD matches.

1. `https://www.alz.washington.edu/`

4. Conclusion

Early diagnosis for dementia subtypes is clinically difficult due to overlapping symptoms. In this work, we demonstrated that machine learning model can outperform clinicians on the mixed AD+LBD subtype detection. However, this conclusion still needs to be validated on a larger test set for this particular task. Also, classification performance of pure LBD and Mix AD+LBD still remains low, indicating the continuous need for clinical and machine learning research.

Table 1: F1 scores of the best model and clinicians

Dementia Subtype	Best Model	Clinicians
Pure AD	0.569 (0.506, 0.633)	0.556 (0.489, 0.624)
Pure LBD	0.000 (0.000, 0.000)	**0.283 (0.100, 0.526)**
Mix AD + LBD	**0.283 (0.183, 0.381)**	0.062 (0.000, 0.133)
Other subtypes	0.523 (0.422, 0.609)	0.584 (0.492, 0.667)

Table 2: Sensitivity of the best model and clinicians

Dementia Subtype	Best Model	Clinicians
Pure AD	0.730 (0.650, 0.805)	0.680 (0.597, 0.761)
Pure LBD	0.000 (0.000, 0.000)	**0.406 (0.000, 0.800)**
Mix AD + LBD	**0.215 (0.130, 0.300)**	0.033 (0.000, 0.075)
Other subtypes	0.464 (0.360, 0.570)	**0.671 (0.567, 0.764)**

Table 3: Specificity of the best model and clinicians

Dementia Subtype	Best Model	Clinicians
Pure AD	0.442 (0.370, 0.511)	0.495 (0.426, 0.563)
Pure LBD	**0.997 (0.990, 1.000)**	0.965 (0.942, 0.983)
Mix AD + LBD	0.864 (0.820, 0.903)	**0.970 (0.946, 0.990)**
Other subtypes	**0.885 (0.837, 0.926)**	0.770 (0.714, 0.822)

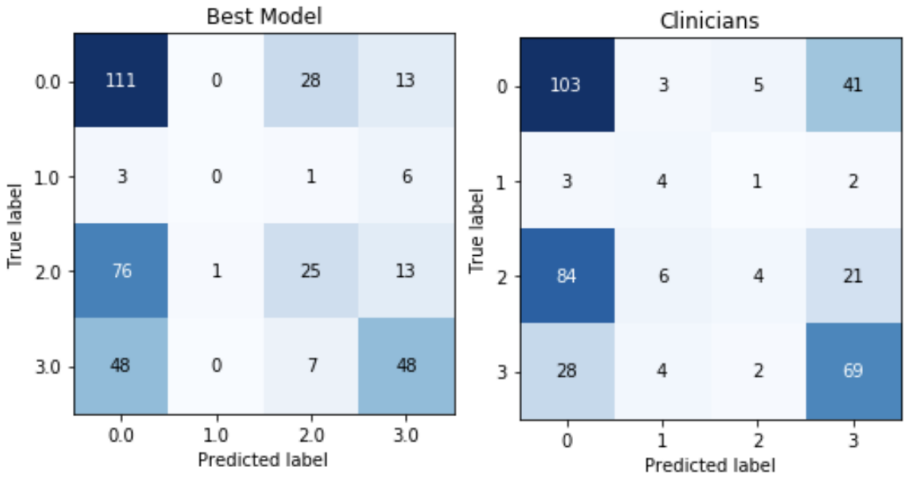

Figure 1: Confusion matrix of the best model and clinicians on the entire test set (0=Pure AD, 1=Pure LBD, 2=Mix AD+LBD, 3=None of these).

References

[1] S. J. Colloby, R. A. Cromarty, L. R. Peraza, K. Johnsen, G. Jóhannesson, L. Bonanni, M. Onofrj, R. Barber, J. T. O'Brien, and J. P. Taylor. Multimodal EEG-MRI in the differential diagnosis of Alzheimer's disease and dementia with Lewy bodies. *Journal of Psychiatric Research*, 78:48–55, 2016.

[2] M. Dauwan, J. J. van der Zande, E. van Dellen, I. E. C. Sommer, P. Scheltens, A. W. Lemstra, and C. J. Stam. Random forest to differentiate dementia with Lewy bodies from Alzheimer's disease. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, 4:99–106, 2016.

[3] A. Katako, P. Shelton, A. L. Goertzen, D. Levin, B. Bybel, M. Aljuaid, H. J. Yoon, D. Y. Kang, S. M. Kim, C. S. Lee, and J. H. Ko. Machine learning identified an Alzheimer's disease-related FDG-PET pattern which is also expressed in Lewy body dementia and Parkinson's disease dementia. *Scientific Reports*, 8(1):1–13, 2018.

[4] A. V. Lebedev, E. Westman, M. K. Beyer, M. G. Kramberger, C. Aguilar, Z. Pirtosek, and D. Aarsland. Multivariate classification of patients with Alzheimer's and dementia with Lewy bodies using high-dimensional cortical thickness measurements: An MRI surface-based morphometric study. *Journal of Neurology*, 260(4):1104–1115, 2013.

[5] H. Lee, G. J. Brekelmans, and G. Roks. The EEG as a diagnostic tool in distinguishing between dementia with Lewy bodies and Alzheimer's disease. *Clinical Neurophysiology*, 126(9):1735–1739, 2015.

[6] K. Oppedal, K. Engan, T. Eftestøl, M. Beyer, and D. Aarsland. Classifying Alzheimer's disease, Lewy body dementia, and normal controls using 3D texture analysis in magnetic resonance images. *Biomedical Signal Processing and Control*, 33:19–29, 2017.

[7] J. J. van der Zande, A. A. Gouw, I. v. Steenoven, P. Scheltens, C. J. Stam, and A. W. Lemstra. EEG characteristics of dementia with Lewy Bodies, Alzheimer's Disease and mixed pathology. *Frontiers in Aging Neuroscience*, 10(JUL):1–10, 2018.

[8] A. Wada, K. Tsuruta, R. Irie, K. Kamagata, T. Maekawa, S. Fujita, S. Koshino, K. Kumamaru, M. Suzuki, A. Nakanishi, M. Hori, and S. Aoki. Differentiating Alzheimer's disease from dementia with lewy bodies using a deep learning technique based on structural brain connectivity. *Magnetic Resonance in Medical Sciences*, 18(3):219–224, 2019.