Fachhochschule Dortmund
University of Applied Sciences and Arts

Machine learning and interpretability methods to investigate Alzheimer's disease

Louise Bloch[1,2,3] and Christoph M. Friedrich[1,2]

[1]Department of Computer Science, University of Applied Sciences and Arts Dortmund (FH Dortmund), 44227 Dortmund, Germany
[2]Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University Hospital Essen, 45122 Essen, Germany
[3]Institute for Artificial Intelligence in Medicine (IKIM), University Hospital Essen, 45122 Essen, Germany

2024-05-15

Alzheimer's Disease (AD)

- Most frequent cause of dementia [1]
- Neurodegeneration starts decades before dementia symptoms occur
- At time of diagnosis, many neurons are irreversibly degenerated
- No cure, only reduction of symptoms [2]
- Early detection important but complex due to heterogenous disease profiles

Motivation

- Machine Learning (ML) to identify complex patterns in high-dimensional data
- Identifying complex patterns that improve the early prediction of AD
- Complex underlying systems require complex models
- Interpretable ML (IML) to explain decisions of black-box models and validated biological plausibility

Research Approach

1. Train ML and Deep-Learning (DL) models to predict AD.
2. Check generalizability during internal and external validation.
3. Use interpretability methods to explain black-box models.
4. Compare explanations of ML- and DL- models to each other.
5. Validate biological plausability of explanations with a ground-truth Voxel-Based Morphometry (VBM) [3].

Data

Table 1: Demographic data, and MRI field strength of the selected subjects, separated by diagnosis groups. For continuous features, mean and standard deviation are given.

Diagnosis	n	Age (years)	Females (%)	1.5 T (%)	3 T (%)
ADNI [4] (for training and internal validation)					
CN	512	74.20 ± 5.82	51.76	44.00	56.00
AD	335	74.95 ± 7.74	44.78	57.00	43.00
Σ	847	74.50 ± 6.66	49.00	49.00	51.00
AIBL [5] (for external validation)					
CN	446	72.53 ± 6.14	56.95	19.06	80.94
AD	71	73.26 ± 7.88	59.15	16.90	83.10
Σ	517	72.63 ± 6.41	57.25	18.76	81.24
OASIS [6] (for external validation)					
CN	704	68.35 ± 9.27	58.66	12.36	87.64
AD	198	75.62 ± 7.92	48.48	10.61	89.39
Σ	902	69.94 ± 9.48	56.43	11.97	88.03

Feature Extraction

- Classical ML models:
  - Volumes of brain-regions extracted from Magnetic-Resonance-Imaging (MRI) scans
  - Normalized by estimated Total Intracranial Volume (eTIV)
- Deep-Learning models:
  - Convolutional Neural Networks (CNNs) trained on skull-stripped 3D-MRI scans

Figure 1: T1-weighted MRI-scan segmented by FreeSurfer v6.0. Adapted from: [7].

Fachhochschule Dortmund
University of Applied Sciences and Arts

Model Training

- ADNI dataset split in 80 % training and 20 % independent test set (stratified)
- Hyperparameter tuning: Grid-search including 5-fold-cross-validation (CV)
- Interpretable-by-design: Decision Trees (DTs), Logisitic Regression (LR)
- Black-Box: Support Vector Machines (SVMs) [8], Random Forest (RF) [9] , eXtreme Gradient Boosting (XGBoost) [10], Light Gradient Boosting (LightGBM) [11]
- Deep Learning (CNNs): DenseNet [12], EfficientNet-B0 [13], Squeeze and Excitation (SE) [14]-ResNet [15], and -ResNeXt [16]
- Platt scaling [17] for model calibration

Interpretability Methods

- Highly correlated features are consolidated into aspects [18]
- All models: SHapley Additive exPlanations (SHAP) [19], Local Interpretable Model-Agnostic Explanations (LIME) [20]
- Classical ML: Permutation-based feature importance
- Deep Learning: Gradient-weighted Class Activation Mapping (GradCAM) [21], GradCAM++ [22]
- Deep Learning explanations summarized for regions

Internal and External Validation

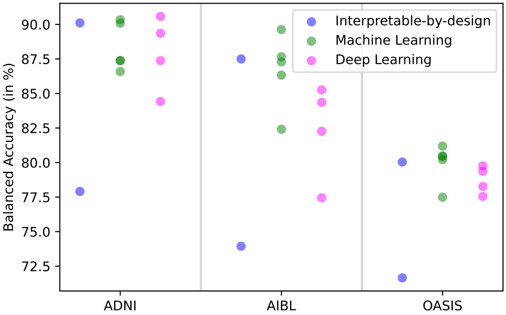

Figure 2: Plot showing performance of the trained ML and DL models.

- Performance of models that are interpretable-by-design show strong differences (performance of DT weak, performance of LR fair)
- Performance of DL models does not outperform classical ML models
- AIBL performances acceptable (generalizability for DL models worse than for classical ML)
- OASIS results acceptable but worse than remaining performances

Explain Classical ML Model Decisions

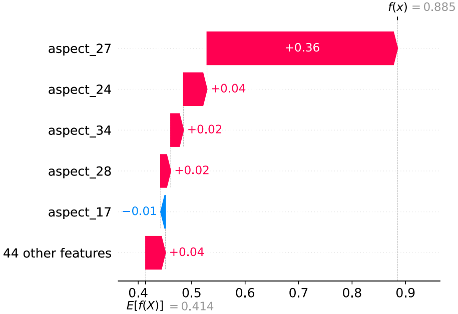

Figure 3: SHapley Additive exPlanations (SHAP) [19] waterfall plot to explain individual decision of a subject with AD for LightGBM model.

- Explain the differences of the individual prediction $(f(x) = 0.414)$ and the average model prediction $(E(f(x)) = 0.885)$ using the model input features
- Feature expressions that increase the AD risk (aspect_27, aspect_24)
- Other feature expressions have a protective influence (aspect_17).

Fachhochschule Dortmund
University of Applied Sciences and Arts

Explain DL Model Decisions

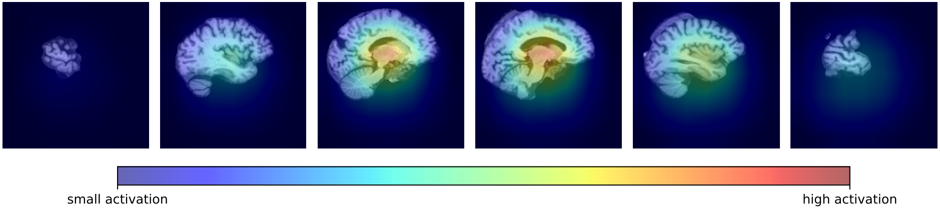

Figure 4: Heatmap showing GradCAM++ results to explain individual decision of a subject with AD for the DenseNet model. Source: [23]

Voxel-based-Morphometry Analysis

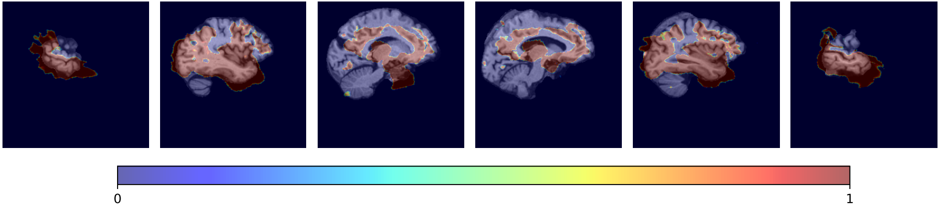

Figure 5: VBM analysis results visualize ground-truth relevant brain regions of a subject with AD. Source: [23]

Comparison to Biologically Plausible Ground Truth

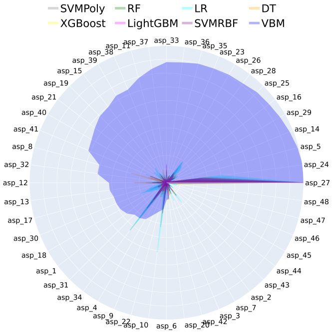

Figure 6: Polar plot to compare classical ML model explanations to VBM ground truth.

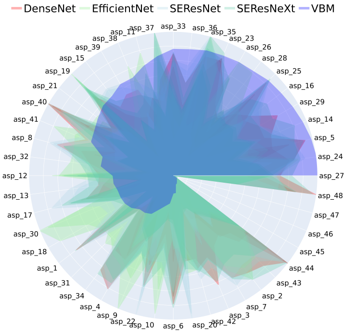

Figure 7: Polar plot to compare Deep-Learning-Model explanations to VBM ground truth.

Future Work

- Check why the localization of the Deep-Learning model explanations is rather unfocused (new information vs. overfitting / underfitting)
- Validation on clinically more relevant research questions (e.g., Mild cognitive impaired subjects, Amyloid-β-positivity, Tau-positivity)
- Use of multimodal input features

Fachhochschule Dortmund
University of Applied Sciences and Arts

References I

[1] C. Patterson. *World Alzheimer report 2018 - The state of the art of dementia research: New frontiers.* https://www.alz.co.uk/research/WorldAlzheimerReport2018.pdf, Accessed: 2021-10-10. London, Great Britain: Alzheimer's Disease International, 2018 (cit. on p. 2).

[2] J. Cao, J. Hou, J. Ping, and D. Cai. "Advances in developing novel therapeutic strategies for Alzheimer's disease". In: *Mol. Neurodegener.* 13.1 (2018). DOI: 10.1186/s13024-018-0299-8 (cit. on p. 2).

[3] J. Ashburner and K. J. Friston. "Voxel-Based Morphometry - The Methods". In: *NeuroImage* 11.6 (2000), pp. 805–821. DOI: 10.1006/nimg.2000.0582 (cit. on p. 4).

[4] R. C. Petersen, P. S. Aisen, L. A. Beckett, M. C. Donohue, A. C. Gamst, D. J. Harvey, et al. "Alzheimer's Disease Neuroimaging Initiative (ADNI)". In: *Neurology* 74.3 (2010), pp. 201–9. DOI: 10.1212/WNL.0b013e3181cb3e25 (cit. on p. 5).

[5] K. A. Ellis, A. I. Bush, D. Darby, D. De Fazio, J. Foster, P. Hudson, et al. "The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: Methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's Disease". In: *International Psychogeriatrics* 21.4 (2009), pp. 672–687 (cit. on p. 5).

[6] P. J. LaMontagne, T. L. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, et al. "OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer Disease". In: *medRxiv* (2019) (cit. on p. 5).

[7] Beijing Normal University. *Beijing Normal University, State Key Laboratory of Cognitive Neuroscience and Learning Enhanced Sample.* Distributed under Creative Commons Attribution-NonCommerical 3.0 Unported License (CC-BY-NC). URL: http://fcon_1000.projects.nitrc.org/indi/retro/BeijingEnhanced.html (visited on 04/17/2024) (cit. on p. 6).

[8] C. Cortes and V. Vapnik. "Support-vector networks". In: *Machine Learning* 20.3 (1995), pp. 273–297. DOI: 10.1007/bf00994018 (cit. on p. 7).

[9] L. Breiman. "Random Forests". In: *Machine Learning* 45.1 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324 (cit. on p. 7).

[10] T. Chen and C. Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD 2016)*. San Francisco, CA, USA: ACM, 2016, pp. 785–794. DOI: 10.1145/2939672.2939785 (cit. on p. 7).

Fachhochschule Dortmund
University of Applied Sciences and Arts

References II

[11] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, et al. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: *Advances in Neural Information Processing Systems (NIPS 2017)*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. Vol. 30. Long Beach, CA, USA: Curran Associates, Inc., 2017, pp. 3146–3154 (cit. on p. 7).

[12] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. "Densely Connected Convolutional Networks". In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*. Honolulu, HI, USA, 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243 (cit. on p. 7).

[13] M. Tan and Q. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by K. Chaudhuri and R. Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, 2019, pp. 6105–6114 (cit. on p. 7).

[14] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu. "Squeeze-and-Excitation Networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)*. Salt Lake City, UT, USA, 2017, pp. 7132–7141. DOI: 10.1109/cvpr.2018.00745 (cit. on p. 7).

[15] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. LasVegas, NV, USA, 2016, pp. 770–778. DOI: 10.1109/cvpr.2016.90 (cit. on p. 7).

[16] S. Xie, R. B. Girshick, P. Dollár, Z. Tu, and K. He. "Aggregated Residual Transformations for Deep Neural Networks". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA, 2017, pp. 5987–5995. DOI: 10.1109/cvpr.2017.634 (cit. on p. 7).

[17] J. C. Platt. "Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods". In: *Advances in large margin classifiers*. MIT Press, 1999, pp. 61–74 (cit. on p. 7).

[18] K. Pekala, K. Woznica, and P. Biecek. *Triplot: Model agnostic measures and visualisations for variable importance in predictive models that take into account the hierarchical correlation structure*. 2021. DOI: 10.48550/arXiv.2104.03403 (cit. on p. 8).

[19] S. M. Lundberg and S.-I. Lee. "A unified approach to interpreting model predictions". In: *Advances in Neural Information Processing Systems (NIPS 2017)*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. Vol. 30. New York, New York, US: Curran Associates, Inc., 2017, pp. 4765–74 (cit. on pp. 8, 10).

Fachhochschule Dortmund
University of Applied Sciences and Arts

References III

[20] M. Ribeiro, S. Singh, and C. Guestrin. "Why should I trust you?: Explaining the predictions of any classifier". In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations (NAACL 2016): 12-17 June 2016; San Diego, California, US*. San Diego, US: Association for Computational Linguistics, 2016, pp. 97–101. DOI: 10.18653/v1/n16-3020 (cit. on p. 8).

[21] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. "Grad-CAM: Visual explanations from deep networks via gradient-based localization". In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV 2017)*. Venice, Italy, 2017, pp. 618–626. DOI: 10.1109/ICCV.2017.74 (cit. on p. 8).

[22] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian. "Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks". In: *IEEE Winter Conference on Applications of Computer Vision (WACV 2018)*. Lake Tahoe, NV, USA, 2018, pp. 839–847. DOI: 10.1109/WACV.2018.00097 (cit. on p. 8).

[23] L. Bloch and C. M. Friedrich. "Systematic comparison of 3D Deep learning and classical machine learning explanations for Alzheimer's Disease detection". In: *Computers in Biology and Medicine* 170 (2024), p. 108029. DOI: 10.1016/j.compbiomed.2024.108029 (cit. on pp. 11, 12).

Fachhochschule Dortmund
University of Applied Sciences and Arts

Closing

Many thanks for your attention!

Please do not hesitate to approach us, given you have any questions!