

Dynamic Machine Learning For Risk Stratification Of MCI-to-AD Conversion Using Longitudinal Multimodal Data:

Mohamed Zahir Elhassan¹, Wen-Xiang Tsai², Carolyn Fredericks¹

Department of Neurology, Yale University, New Haven, CT, USA¹, Institute of Brain Science, National Yang Ming Chiao Tung University, Taipei, Taiwan²

Background

- Alzheimer's disease (AD) pathology begins years before dementia, but **not all individuals with MCI progress to AD**.
- Accurate prediction of **MCI-to-AD conversion** could improve patient counseling, early intervention, and clinical trial enrichment.
- Longitudinal clinical datasets are **multimodal, irregularly sampled, and contain substantial missing data**, limiting conventional prediction models.
- Traditional machine learning often relies on static baseline measurements and cannot fully capture temporal disease progression.
- We evaluated whether **trajectory-aware machine learning and deep learning** improve individualized prediction of near-term conversion.

Methods

Cohort:

- 204 biomarker-positive MCI participants from ADNI
- 61 converted to AD dementia within 3 years

Feature Sets:

- Cognitive assessments
- MRI volumetric measures
- Amyloid and tau biomarkers
- Demographics
- Combined multimodal dataset

Models Compared:

- Logistic Regression
- Random Forest
- Time-varying Cox Regression
- Trajectory-enhanced models
- Deep longitudinal neural network
 - GRU-D
 - Neural ODE
 - Multi-head attention
 - Dynamic proportional hazards

Evaluation:

- Nested cross-validation
- AUROC
- AUPRC
- Balanced Accuracy
- F1 Score

Results

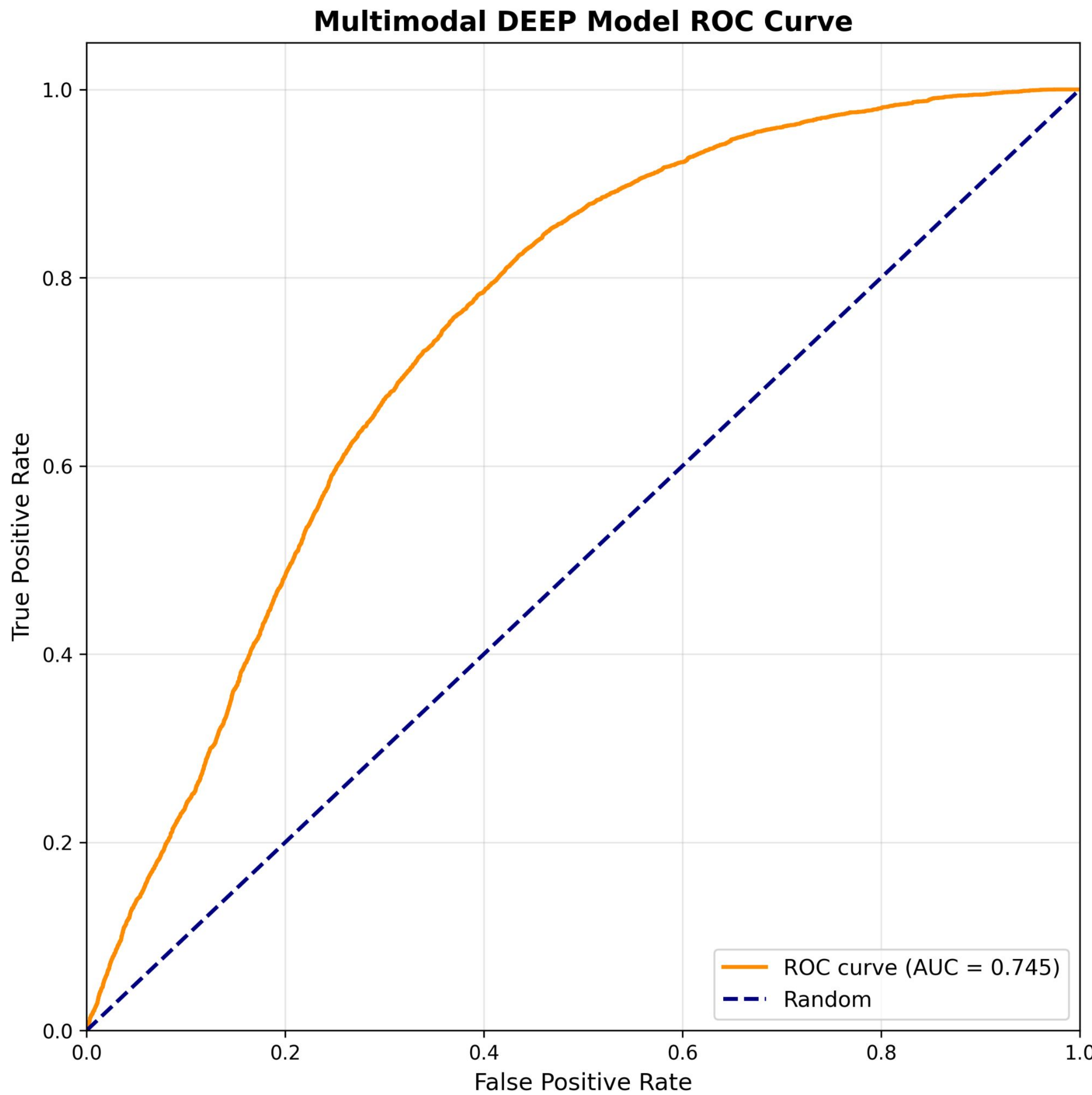


Table 1: Model Performance at Baseline by Feature Configuration

Configuration	Model	AUROC	AUPRC	Bal. Acc.	F1-opt
Cognitive + Demographics	Deep Learning	0.736 ± 0.100	0.632 ± 0.102	0.640 ± 0.080	0.538 ± 0.109
	LogReg	0.754 ± 0.086	0.645 ± 0.085	0.689 ± 0.103	0.602 ± 0.129
	RF	0.772 ± 0.090	0.683 ± 0.121	0.691 ± 0.086	0.606 ± 0.113
	LogReg-Traj	0.757 ± 0.080	0.643 ± 0.083	0.689 ± 0.103	0.613 ± 0.139
	RF-Traj	0.772 ± 0.088	0.677 ± 0.128	0.698 ± 0.109	0.604 ± 0.126
	Cox-Traj	0.761 ± 0.067	0.650 ± 0.099	0.621 ± 0.052	0.605 ± 0.032
Pathology + Demographics	Deep Learning	0.789 ± 0.083	0.698 ± 0.103	0.709 ± 0.114	0.686 ± 0.097
	LogReg	0.781 ± 0.102	0.688 ± 0.140	0.711 ± 0.110	0.632 ± 0.119
	RF	0.789 ± 0.088	0.699 ± 0.102	0.694 ± 0.109	0.592 ± 0.107
	LogReg-Traj	0.780 ± 0.102	0.688 ± 0.141	0.707 ± 0.114	0.629 ± 0.138
	RF-Traj	0.792 ± 0.076	0.704 ± 0.098	0.701 ± 0.110	0.631 ± 0.126
	Cox-Traj	0.739 ± 0.107	0.632 ± 0.157	0.628 ± 0.060	0.600 ± 0.049
Volumes + Demographics	Deep Learning	0.760 ± 0.083	0.644 ± 0.113	0.669 ± 0.079	0.653 ± 0.081
	LogReg	0.771 ± 0.074	0.673 ± 0.115	0.699 ± 0.086	0.656 ± 0.078
	RF	0.782 ± 0.052	0.697 ± 0.049	0.651 ± 0.073	0.568 ± 0.076
	LogReg-Traj	0.771 ± 0.073	0.675 ± 0.113	0.699 ± 0.086	0.639 ± 0.094
	RF-Traj	0.774 ± 0.049	0.681 ± 0.050	0.652 ± 0.062	0.572 ± 0.085
	Cox-Traj	0.771 ± 0.060	0.675 ± 0.091	0.574 ± 0.018	0.577 ± 0.014
All Modalities (Multimodal)	Deep Learning	0.754 ± 0.081	0.657 ± 0.101	0.655 ± 0.116	0.603 ± 0.104
	LogReg	0.772 ± 0.082	0.651 ± 0.116	0.694 ± 0.097	0.625 ± 0.154
	RF	0.788 ± 0.059	0.714 ± 0.067	0.684 ± 0.063	0.571 ± 0.073
	LogReg-Traj	0.768 ± 0.088	0.643 ± 0.126	0.701 ± 0.103	0.618 ± 0.135
	RF-Traj	0.790 ± 0.058	0.718 ± 0.077	0.719 ± 0.053	0.613 ± 0.051
	Cox-Traj	0.734 ± 0.102	0.617 ± 0.147	0.617 ± 0.050	0.591 ± 0.034

Objectives/Aims

Primary Aim:

- Develop a longitudinal machine learning framework to predict **3-year conversion from MCI to Alzheimer's disease**.

Secondary Aims:

- Compare classical machine learning, survival analysis, and deep learning approaches.
- Determine whether incorporating **longitudinal trajectories** improves prediction beyond baseline measurements.
- Evaluate the contribution of cognitive, imaging, pathological, and multimodal feature sets.

Conclusions

- Trajectory-based models modestly improved prediction** over static baseline approaches.
- Random Forest with trajectory features achieved the strongest overall performance** across multimodal data (AUROC 0.790, AUPRC 0.718, Balanced Accuracy 0.719).
- Deep learning achieved competitive performance but **did not consistently outperform simpler machine learning models**.
- Increased model complexity alone **did not translate into superior discrimination** in this moderate-sized longitudinal cohort.
- Longitudinal information appears more valuable than algorithmic complexity for short-term AD risk prediction.

Future Directions

- Evaluate models at **later clinical landmarks** (6-, 12-, and 24-month follow-up).
- Expand validation in larger, independent multicenter cohorts.
- Incorporate **additional modalities** (genetics, plasma biomarkers, PET imaging).
- Improve model interpretability using explainable AI methods.
- Develop clinically deployable tools for individualized AD risk stratification.
- Translation into clinician-facing **decision support tools** to facilitate individualized prognosis and clinical trial enrollment.

Mohamed Z. Elhassan
Stanford Class of 2025
Biomedical Computation (B.S.)
mohamedzahir20@gmail.com

Acknowledgements: We thank the Alzheimer's Disease Neuroimaging Initiative (ADNI) investigators and participants for providing the data used in this study. We also thank members of the Yale Neurology Department for their guidance and support. Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

References: 1 Petersen RC, et al. J Intern Med. 2014. 2 Li D, et al. Brain Sciences. 2021. 3 Pontecorvo MJ, et al. JAMA Neurol. 2023. 4 Che Z, et al. Sci Rep. 2018. 5 Chen RTQ, et al. NeurolPS. 2018. 6 Vaswani A, et al. NeurolPS. 2017. 7 Alzheimer's Disease Neuroimaging Initiative (ADNI).