



Machine Learning–Based Multi-Trait Polygenic Risk Modeling Improves Alzheimer’s Disease Classification

Saira Kiran¹, Jingchun Chen^{1,2,3}

¹Interdisciplinary Neuroscience Program, UNLV; ² Nevada Institute of Personalized Medicine, UNLV; ³School of Integrated Health Sciences, UNLV

Poster Code: aaic2026.ed31ac6

Alzheimer's disease has a complex genetic architecture, and conventional polygenic risk scores (PRS) provide only limited predictive power. This study evaluates whether integrating multi-trait PRS with demographic and genetic covariates using machine learning improves Alzheimer's disease classification.

INTRODUCTION	RESULTS							DISCUSSION																																																																																																																																																																														
• Background: Polygenic risk scores (PRS) derived from genome-wide association studies (GWAS) provide a promising approach for estimating genetic susceptibility to Alzheimer's disease (AD). However, PRS alone captures only a fraction of the disease's complex genetic architecture. • The Problem: Conventional single-trait PRS models have limited predictive accuracy because they do not account for the combined effects of multiple genetic traits or important demographic and genetic risk factors. • Objective: We investigated whether integrating multi-trait PRS with age, sex, APOE ε4 status, and population structure using machine learning improves Alzheimer's disease classification compared with PRS-only models.	• Data set: 2,651 AD cases and 2,768 controls. • PRS Traits: 21 total (1 AD-specific, 20 non-AD) • Covariates: Age, Sex, APOE ε4 carrier status (rs429358, rs7412), and top 10 Principal Components (PCs). <table><tr><th colspan="8">Table 1. Performance from Feature Set 1 (PRS Only)</th></tr><tr><td rowspan="6">AD</td><th>Model</th><th>AUC</th><th>Accuracy</th><th>Precision</th><th>Recall</th><th>F1_Score</th><th>Sensitivity</th></tr><tr><td>CatBoost</td><td>0.651</td><td>0.603</td><td>0.599</td><td>0.57</td><td>0.584</td><td>0.57</td></tr><tr><td>DNN</td><td>0.669</td><td>0.629</td><td>0.617</td><td>0.638</td><td>0.627</td><td>0.638</td></tr><tr><td>LR</td><td>0.675</td><td>0.64</td><td>0.634</td><td>0.626</td><td>0.63</td><td>0.626</td></tr><tr><td>Random Forest</td><td>0.664</td><td>0.607</td><td>0.594</td><td>0.619</td><td>0.606</td><td>0.619</td></tr><tr><td>XGBoost</td><td>0.61</td><td>0.57</td><td>0.565</td><td>0.525</td><td>0.544</td><td>0.525</td></tr><tr><td rowspan="6">Non-AD</td><td>CatBoost</td><td>0.604</td><td>0.552</td><td>0.544</td><td>0.509</td><td>0.526</td><td>0.509</td></tr><tr><td>DNN</td><td>0.621</td><td>0.578</td><td>0.577</td><td>0.509</td><td>0.541</td><td>0.509</td></tr><tr><td>LR</td><td>0.627</td><td>0.583</td><td>0.578</td><td>0.543</td><td>0.56</td><td>0.543</td></tr><tr><td>Random Forest</td><td>0.62</td><td>0.568</td><td>0.561</td><td>0.54</td><td>0.55</td><td>0.540</td></tr><tr><td>XGBoost</td><td>0.594</td><td>0.57</td><td>0.562</td><td>0.547</td><td>0.555</td><td>0.547</td></tr></table> <table><tr><th colspan="8">Table 2. Performance from Feature set 10 (PRS + Covariates)</th></tr><tr><td rowspan="6">AD</td><th>Model</th><th>AUC</th><th>Accuracy</th><th>Precision</th><th>Recall</th><th>F1_Score</th><th>Sensitivity</th></tr><tr><td>CatBoost</td><td>0.818</td><td>0.755</td><td>0.758</td><td>0.732</td><td>0.745</td><td>0.732</td></tr><tr><td>DNN</td><td>0.799</td><td>0.731</td><td>0.73</td><td>0.713</td><td>0.721</td><td>0.713</td></tr><tr><td>LR</td><td>0.785</td><td>0.72</td><td>0.718</td><td>0.702</td><td>0.71</td><td>0.702</td></tr><tr><td>Random Forest</td><td>0.793</td><td>0.731</td><td>0.721</td><td>0.732</td><td>0.727</td><td>0.732</td></tr><tr><td>XGBoost</td><td>0.776</td><td>0.69</td><td>0.673</td><td>0.713</td><td>0.692</td><td>0.713</td></tr><tr><td rowspan="6">Non-AD</td><td>CatBoost</td><td>0.816</td><td>0.745</td><td>0.747</td><td>0.725</td><td>0.736</td><td>0.725</td></tr><tr><td>DNN</td><td>0.798</td><td>0.71</td><td>0.714</td><td>0.679</td><td>0.696</td><td>0.679</td></tr><tr><td>LR</td><td>0.784</td><td>0.723</td><td>0.72</td><td>0.709</td><td>0.715</td><td>0.709</td></tr><tr><td>Random Forest</td><td>0.776</td><td>0.72</td><td>0.713</td><td>0.713</td><td>0.713</td><td>0.713</td></tr><tr><td>XGBoost</td><td>0.77</td><td>0.712</td><td>0.701</td><td>0.717</td><td>0.709</td><td>0.717</td></tr></table>							Table 1. Performance from Feature Set 1 (PRS Only)								AD	Model	AUC	Accuracy	Precision	Recall	F1_Score	Sensitivity	CatBoost	0.651	0.603	0.599	0.57	0.584	0.57	DNN	0.669	0.629	0.617	0.638	0.627	0.638	LR	0.675	0.64	0.634	0.626	0.63	0.626	Random Forest	0.664	0.607	0.594	0.619	0.606	0.619	XGBoost	0.61	0.57	0.565	0.525	0.544	0.525	Non-AD	CatBoost	0.604	0.552	0.544	0.509	0.526	0.509	DNN	0.621	0.578	0.577	0.509	0.541	0.509	LR	0.627	0.583	0.578	0.543	0.56	0.543	Random Forest	0.62	0.568	0.561	0.54	0.55	0.540	XGBoost	0.594	0.57	0.562	0.547	0.555	0.547	Table 2. Performance from Feature set 10 (PRS + Covariates)								AD	Model	AUC	Accuracy	Precision	Recall	F1_Score	Sensitivity	CatBoost	0.818	0.755	0.758	0.732	0.745	0.732	DNN	0.799	0.731	0.73	0.713	0.721	0.713	LR	0.785	0.72	0.718	0.702	0.71	0.702	Random Forest	0.793	0.731	0.721	0.732	0.727	0.732	XGBoost	0.776	0.69	0.673	0.713	0.692	0.713	Non-AD	CatBoost	0.816	0.745	0.747	0.725	0.736	0.725	DNN	0.798	0.71	0.714	0.679	0.696	0.679	LR	0.784	0.723	0.72	0.709	0.715	0.709	Random Forest	0.776	0.72	0.713	0.713	0.713	0.713	XGBoost	0.77	0.712	0.701	0.717	0.709	0.717	• Multi-trait PRS improves AD prediction compared to single-trait PRS. • Integrating genetic traits with covariates enhances model performance. • Non-AD traits also contribute to AD risk prediction. • Machine learning effectively integrates high-dimensional genetic data.
Table 1. Performance from Feature Set 1 (PRS Only)																																																																																																																																																																																						
AD	Model	AUC	Accuracy	Precision	Recall	F1_Score	Sensitivity																																																																																																																																																																															
	CatBoost	0.651	0.603	0.599	0.57	0.584	0.57																																																																																																																																																																															
	DNN	0.669	0.629	0.617	0.638	0.627	0.638																																																																																																																																																																															
	LR	0.675	0.64	0.634	0.626	0.63	0.626																																																																																																																																																																															
	Random Forest	0.664	0.607	0.594	0.619	0.606	0.619																																																																																																																																																																															
	XGBoost	0.61	0.57	0.565	0.525	0.544	0.525																																																																																																																																																																															
Non-AD	CatBoost	0.604	0.552	0.544	0.509	0.526	0.509																																																																																																																																																																															
	DNN	0.621	0.578	0.577	0.509	0.541	0.509																																																																																																																																																																															
	LR	0.627	0.583	0.578	0.543	0.56	0.543																																																																																																																																																																															
	Random Forest	0.62	0.568	0.561	0.54	0.55	0.540																																																																																																																																																																															
	XGBoost	0.594	0.57	0.562	0.547	0.555	0.547																																																																																																																																																																															
	Table 2. Performance from Feature set 10 (PRS + Covariates)																																																																																																																																																																																					
AD	Model	AUC	Accuracy	Precision	Recall	F1_Score	Sensitivity																																																																																																																																																																															
	CatBoost	0.818	0.755	0.758	0.732	0.745	0.732																																																																																																																																																																															
	DNN	0.799	0.731	0.73	0.713	0.721	0.713																																																																																																																																																																															
	LR	0.785	0.72	0.718	0.702	0.71	0.702																																																																																																																																																																															
	Random Forest	0.793	0.731	0.721	0.732	0.727	0.732																																																																																																																																																																															
	XGBoost	0.776	0.69	0.673	0.713	0.692	0.713																																																																																																																																																																															
Non-AD	CatBoost	0.816	0.745	0.747	0.725	0.736	0.725																																																																																																																																																																															
	DNN	0.798	0.71	0.714	0.679	0.696	0.679																																																																																																																																																																															
	LR	0.784	0.723	0.72	0.709	0.715	0.709																																																																																																																																																																															
	Random Forest	0.776	0.72	0.713	0.713	0.713	0.713																																																																																																																																																																															
	XGBoost	0.77	0.712	0.701	0.717	0.709	0.717																																																																																																																																																																															
	CONCLUSION																																																																																																																																																																																					
• Multi-trait polygenic risk score (PRS) modeling improves the prediction of Alzheimer’s disease compared to traditional single-trait PRS approaches. Integrating multiple genetic traits with key covariates such as age, sex, and APOE status enhances model performance and provides a more comprehensive representation of genetic risk. These findings highlight the value of combining diverse genetic information with machine learning methods to improve Alzheimer’s disease risk prediction and support future precision medicine strategies.																																																																																																																																																																																						
References																																																																																																																																																																																						
• He Z, et al. The current status, trends, and challenges of Alzheimer’s disease and other dementias in Asia (1990–2036). Front Public Health. 2025 Jun 10;13. • Dementia. Available from: https://www.who.int/news-room/fact-sheets/detail/dementia • Alzheimer’s Facts and Figures Report Alzheimer’s Association. Available from: https://www.alz.org/alzheimers-dementia/facts-figures • From Amyloid to Synaptic Dysfunction: Biomarker-Driven Insights into Alzheimer’s Disease. Available from: https://www.mdpi.com/1467-3045/47/8/580 • Andrews SJ, et al. The complex genetic architecture of Alzheimer’s disease: novel insights and future directions. eBioMedicine. 2023 Apr 1;90:104511.																																																																																																																																																																																						
CONTACT																																																																																																																																																																																						
Email: kiran@unlv.evada.edu																																																																																																																																																																																						
 ALZHEIMER'S ASSOCIATION AAIC > 26																																																																																																																																																																																						

