Monte Carlo Dropout for Uncertainty-Aware Alzheimer's Disease Classification Using Transformer Models on Whole Genome Sequencing

Taeho Jo*, Eun Hye Lee
Department of Radiology and Imaging Sciences, Indiana Alzheimer's Disease Research Center, Center for Neuroimaging, Indiana University School of Medicine, Indianapolis, IN, USA

BACKGROUND

Alzheimer's disease (AD) is a complex neurodegenerative disorder with a polygenic architecture where dozens of genetic variants each confer modest effects. While deep learning approaches have improved genomic prediction accuracy, the reliability of predictions remains a critical concern in clinical applications. Black-box models often exhibit overconfidence even when incorrect, potentially leading to inappropriate clinical interventions. This study introduces an uncertainty-aware classification framework that leverages Monte Carlo Dropout to quantify predictive uncertainty in AD genomic classification. By maintaining dropout active during inference and performing multiple stochastic forward passes, we estimate per-sample variance to identify cases requiring additional clinical scrutiny.

METHODS

We analyzed Whole Genome Sequencing (WGS) data from 1,050 participants (443 cognitively normal, 607 AD dementia) from ADNI and ADSP-FUS1-ADNI-WGS-2. After quality control, approximately 100,000 common variants per individual were retained. The genomic data was segmented into non-overlapping windows of 100 SNPs, which were processed through a transformer architecture with two encoder layers, four attention heads, and feed-forward dimension of 512. Monte Carlo Dropout (rate=0.2) remained active during both training and inference, enabling uncertainty estimation through predictive variance. In parallel, a Random Forest classifier was trained on flattened genomic data. The ensemble combined both models using a learned weighting factor (α=0.6) optimized through 5-fold cross-validation. A variance threshold was established to categorize predictions as "uncertain" or "certain" based on ensemble variance estimates.

Diagnosis Group	Sample Size	Mean Age (SD)	% Males / Females	APOE ε2/ε2	APOE ε2/ε3	APOE ε2/ε4	APOE ε3/ε3	APOE ε3/ε4	APOE ε4/ε4	APOE ε4 Carriers (%)
Cognitive normal (CN)	443	72.63 (6.31)	53.05 / 46.95	2	57	7	257	110	10	27.09
Alzheimer's disease (AD)	607	73.97 (7.25)	42.01 / 57.99	1	19	14	194	278	101	62.44

Table 1. Demographics and APOE genotype distribution by diagnosis group.

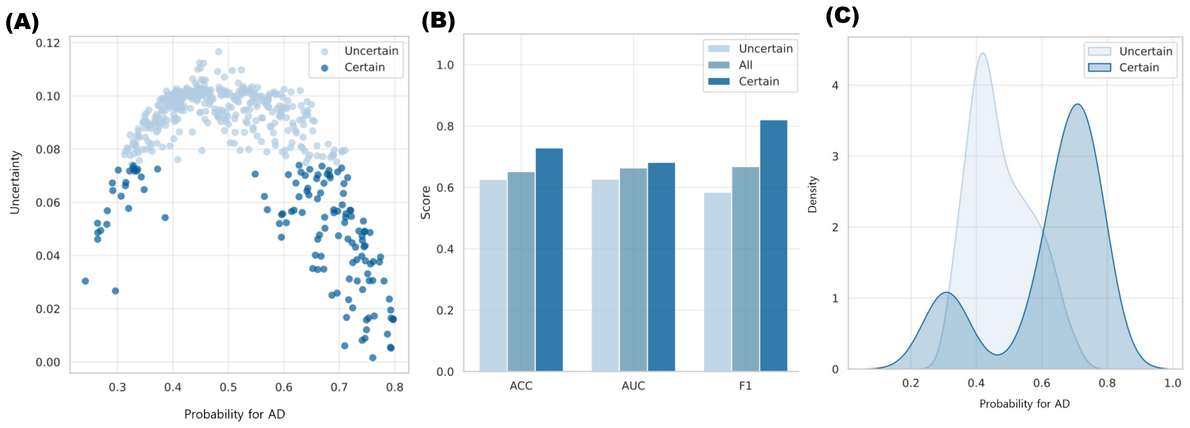

Figure 1. Uncertainty stratification and performance. (A) Probability vs. uncertainty scatter plot. (B) Performance metrics by certainty group. (C) Probability distributions for uncertain/certain samples.

RESULTS

The uncertainty-aware framework achieved an overall accuracy of 0.6514, AUC of 0.6636, and F1-score of 0.6679 on 525 test samples. Variance-based stratification revealed marked performance differences: the certain group (24.6% of samples) achieved accuracy of 0.7287, AUC of 0.6816, and F1-score of 0.8205, representing improvements of 10.24%, 5.48%, and 23.62% respectively over the uncertain group. The uncertain samples (75.4%) showed clustering around intermediate probability values (0.3-0.5), while certain samples appeared at more definitive ranges (<0.3 or >0.6). Among AD patients, 62.44% were APOE ε4 carriers compared to 27.09% in controls, confirming known genetic associations. The optimal ensemble weight of α=0.6 outperformed both individual models, with the transformer alone (α=1.0) yielding the lowest AUC of 0.6283.

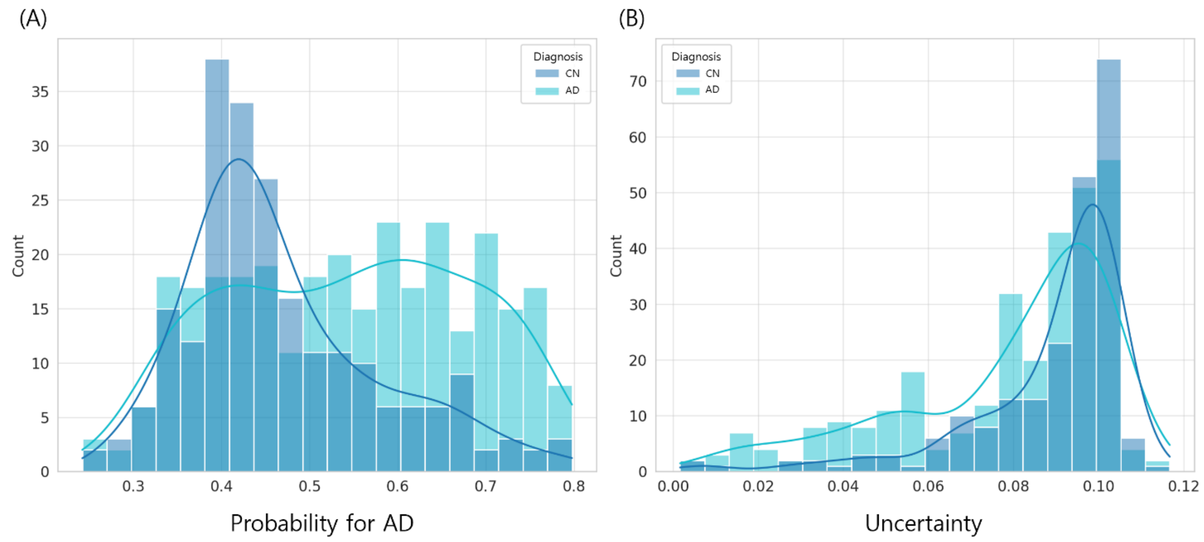

Figure 2. (A) Distribution of predicted probabilities: CN peaks near 0.4, AD shifts toward 0.6. (B) Model uncertainty distribution: CN peaks at 0.10, AD slightly lower, with substantial overlap between groups.

CONCLUSION

This study demonstrates that uncertainty quantification through Monte Carlo Dropout can effectively identify genomic predictions requiring additional clinical validation. By flagging high-variance predictions, the framework achieved substantial performance improvements on confidently classified samples while highlighting ambiguous cases for further diagnostic attention. This approach addresses a critical gap in genomic risk assessment where misclassification can have significant clinical consequences. The integration of transformer-based sequence modeling with traditional machine learning through uncertainty-aware ensemble methods provides a robust framework for complex genomic analysis. Future directions include expansion to multi-ethnic populations, incorporation of rare variants, and integration with other biomarkers to enhance both predictive accuracy and clinical utility in AD risk assessment.

Funding
NIH Grants: P30AG010133, P30AG072976, R01AG019771, U01AG068057, U19AG074879; Alzheimer's Association: AARG-22-974053