

EHR 및 코호트 기반 임상 데이터를 활용한 TabTransformer 기반 알츠하이머병 단계 분류 모델

송정아*¹, 윤태현**², 하지환*²

TabTransformer-based Model for Alzheimer's Disease Classification using EHR and Cohort-based Clinical Data

Jung-a Song*¹, Taehyeon Yun**², and Jihwan Ha*²

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2026-25470406)
이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (RS-2023-00242528)

요 약

경도인지장애(MCI)와 알츠하이머병(AD)은 기억력과 일상생활 수행 능력을 저하시키는 대표적 신경퇴행성 질환으로, 고령화의 진행과 함께 유병률이 증가하고 있다. 이에 따라 조기 진단과 정확한 질병 단계 판별을 위한 기계 학습 기반 자동 분류 연구가 활발히 이루어지고 있으나, 기존 연구는 MRI 등 고비용 뇌 영상 데이터에 주로 의존하여 임상적 접근성과 활용성에 한계를 보인다. 본 연구는 전자의무기록(EHR)과 코호트 기반 구조화된 임상 테이블 데이터를 활용하여 TabTransformer 기반 딥러닝 단계 분류 모델을 설계하고, 정상군, 경도인지장애군, 알츠하이머병군 간 집단 조합별 이진 분류를 수행하였다. 층화된 10겹 교차 검증 결과, 모든 과제에서 90% 이상의 정확도를 달성하였으며 기본 다층 퍼셉트론(MLP) 대비 향상된 성능을 보였다. 또한 순열 중요도 분석을 통해 CDR과 MMSE이 핵심 변수로 작용함을 확인하였다.

Abstract

Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) are prevalent neurodegenerative disorders that impair memory and daily functioning, with incidence increasing alongside global population aging. As early diagnosis and accurate stage classification have gained importance, machine learning-based approaches have been widely explored. However, many existing methods depend on costly neuroimaging modalities such as Magnetic Resonance Imaging (MRI), limiting clinical applicability. In this study, we propose a TabTransformer-based deep learning framework that leverages structured clinical tabular data derived from Electronic Health Records (EHR) and cohort information. Binary classification was conducted among cognitively normal, MCI, and AD groups. Using stratified 10-fold cross-validation, the proposed model achieved over 90% accuracy across all tasks and outperformed a baseline Multilayer Perceptron (MLP). Permutation importance analysis further identified Clinical Dementia Rating (CDR) and Mini-Mental State Examination (MMSE) scores as the most influential features.

Keywords

deep learning, TabTransformer, Alzheimer's disease, EHR, cohort

* 부경대학교 빅데이터융합전공(*² 교신저자)

- ORCID¹: <https://orcid.org/0009-0005-2372-7733>

- ORCID²: <https://orcid.org/0000-0002-6086-5693>

** 부경대학교 데이터공학과

- ORCID: <https://orcid.org/0009-0009-3548-7430>

· Received: Mar. 11, 2026, Revised: Apr. 22, 2026, Accepted: Apr. 25, 2026

· Corresponding Author: Jihwan Ha

Major of Big Data Convergence, Division of Data Information Science,
Pukyong National University, Busan 48513, Korea

Tel.: +82-51-629-4614, Email: jhha@pknu.ac.kr

1. 서 론

치매의 가장 흔한 형태는 알츠하이머병(AD, Alzheimer's Disease)으로, 전체 치매 환자의 약 60~70%를 차지하는 것으로 보고된다[1]. AD는 단순한 생리적 노화와 구별되는 진행성 신경퇴행성 질환으로, 초기에는 주로 기억력 저하가 나타나며, 질환이 진행됨에 따라 언어 기능, 실행 기능, 판단력 등 다양한 인지 영역 전반에 걸친 기능 저하가 동반된다. 이에 따라 일상생활 수행 능력이 점진적으로 감소하고, 우울, 불안, 공격성 등 다양한 신경정신과적 증상이 함께 나타날 수 있다[2]-[4]. 한편, 정상 인지 상태와 AD 사이의 이행 단계로 경도 인지 장애(MCI, Mild Cognitive Impairment)가 정의되며, 이는 객관적인 인지 기능 저하가 관찰되지만 기본적인 일상생활 수행 능력은 비교적 유지되는 임상적 상태를 의미한다[5]-[9]. MCI는 치매 진단 기준에는 해당하지 않으나, 일부 환자에서 AD로 진행되는 경향이 보고되고 있어[6][10], AD의 고위험 전 단계로 간주되며 조기 진단 및 예후 예측의 측면에서 중요한 임상적 의의를 갖는다.

인간 평균 수명이 증가하면서 인지 장애 환자는 중장년층 이상의 세대에서 흔히 볼 수 있게 되었다. 미국에서는 2014년 AD 환자가 약 520만 명에 달했으며, 2050년에는 약 1,380만 명에 이를 것으로 예측한다[11]. 대한치매학회[12]는 한국의 치매 유병률은 2003년에 10만 명당 약 178.11명, 2015년 약 5,319.01명으로 급증했으며, 치매 발생률은 2003년 126.41명에서 2015년 2,218.25명으로 급증했다고 보고했다. 급격히 증가하는 AD를 예방하기 위해서는 조기 진단과 정확한 판별이 중요하다.

AD 진단은 뇌 영상 촬영 또는 검사 도구를 이용해 이뤄진다[13]. 뇌 영상 촬영 기법으로는 MRI, CT, PET 등이 있다. 특히 MRI 촬영은 비침습적(손상이 가해지지 않는)이고 방사능 노출이 적어 다른 방법에 비해 촬영이 편리하고 환자 친화적이다[14]. 임상 검사 도구로는 치매척도검사(CDR, Clinical Dementia Rating), 간이인지기능검사(MMSE, Mini-Mental State Examination), 베타-아밀로이드 단백질 및 타우 단백질 검사, 아포지단백 E(APOE,

Apolipoprotein E) 유전자 검사가 있다[15]. APOE는 2형, 3형, 4형(이하 APOE4)의 세 가지 변이 형태가 있으며, APOE4가 치매를 발병하는 데 가장 위험한 요인으로 알려진 바 있다[16]-[21]. J. Fortea et al. [17]는 APOE4 동형접합 집단은 증상 발현의 평균 나이가 약 65세인 반면, 비보유자 집단은 약 73세로, 해당 집단은 약 60-80%의 치매 발병 위험을 지닌다고 하였다. 하지만 AD 검사 시, MRI 촬영을 바로 시도하거나 단일로 진단 결과를 내리는 데 사용하지는 않는다[22]. AD 치매로의 진행에 대한 조기 표지자로서 민감도는 높으나 특이도는 낮기 때문에 MCI에서 AD로의 진행에 관한 조기 진단 도구로써 유용할 뿐, 단일 검사 도구로 사용하기는 권고하지 않는다[12]. 또한 MRI 설치와 유지에 있어 상당한 제약을 받기 때문에 모든 의료기관 및 환자들의 활용에 있어서 제한적이다[23][24]. MRI 장비는 10톤이 넘으므로 장비를 설치할 공간과 전력을 할당하기가 여간 쉽지 않으며, 구매 가격이 자기장 1테슬라(자기장의 세기 단위)당 약 100만 달러에 달할 정도로 고가이다[25]. 그러므로 군 병원, 중환자실, 재난 지역, 인구 소멸 지역 등 접근이 열악한 지역에는 공급이 어렵고, 정기적으로 검사를 받아야 하는 노년층에게는 막대한 의료 비용이 부담된다[26].

이러한 한계점으로 인해, 영상 데이터의 의존도를 낮추고 임상 검사 결과와 기본 인구학적·신체 정보와 같이 비교적 접근과 획득이 용이한 테이블 데이터만으로도 질병 상태를 분류할 수 있는 방법론이 필요하다. 최근 딥러닝(Deep learning) 기반 방법론은 대규모 데이터 학습과 표현 학습 능력을 바탕으로 AD 연구뿐 아니라 의료영상 및 생물정보학 영역의 다양한 데이터 분석에서 표준적 접근으로 자리 잡고 있다[27]-[41].

이에 본 연구는 ADNI(Alzheimer's Disease Neuroimaging Initiative) 코호트 기반 테이블 데이터를 활용하여, 어텐션 메커니즘을 내재한 딥러닝 모델인 TabTransformer의 적용 가능성을 체계적으로 검증하고자 하였다. 기존의 전통적 머신러닝 기법 및 일반적인 딥러닝 기반 단계 분류 모델과 비교 분석을 수행한 결과, 제안 모델은 모든 분류 과제에서 90% 이상의 정확도를 기록하였다. 특히 임상적

으로 구분이 가장 어려운 MCI와 AD 간 분류 과제에서 선행 연구 대비 향상된 성능을 나타내어, 범주형 및 연속형 특성이 혼재된 테이블 데이터 환경에서 어텐션 기반 표현 학습 구조의 우수성을 확인하였다. 더불어 변수 중요도를 정량적으로 분석함으로써 모델의 의사결정 과정을 해석 가능하도록 제시하였으며, 이를 통해 임상적 적용 가능성에 대한 논의를 확장하였다.

II. 관련 연구

2.1 선행 연구

기존 연구는 뇌 영상 기반 질병 상태를 분류하는 연구가 활발하였다. D. Yang et al.[42] fMRI 데이터를 기반으로 CNN, Bi-LSTM(양방향 장단기 메모리), 멀티 헤드 어텐션을 통합한 모델을 제안하여 AD과 정상 상태를 구분하였다. A. Alorf, M.U.G. Khan[43]은 rs-fMRI 데이터로 CN(정상), SMC(Significant Memory Concern), EMCI(Early MCI), MCI, LMCI(Late MCI), AD 6개의 상태를 각각 SSAE 모델과 BC-GCN 모델을 사용하여 분류했다. J. Noh et al.[44]는 rs-fMRI 데이터로 3D-CNN-LSTM 모델로 CN, EMCI, LMCI, AD를 분류했다. N. Chauhan, B. Choi[45]는 다층 ELM 모델을 이용하여 383개의 rs-fMRI 데이터로 AD vs. CN에서 95%, AD vs. MCI에서 81%, CN vs. MCI에서 94%의 정확도를 보였다. 그러나 영상 모달리티(Modalities)는 고차원성과 고비용, 낮은 접근성 때문에 다양한 발병 패턴을 학습하여 실제 의료 현장에 일반화하기 어렵다.

F. Abuhantash et al.[46]는 이미지 기반 연구의 한계점에 주목하여 EHR과 코호트 데이터를 활용한 테이블 단일 모달리티로 AD 분류를 수행했다. GNN(Graph Neural Network)을 변형한 GCN(Graph Convolutional Network), CCN(ChebConv Network), GAT(Graph Attention Network), GraphSAGE 네 가지 모델로 CN, MCI, 성능을 검증했으며 CCN이 정확도 99%(AD vs. CN), 93%(AD vs. MCI), 94%(CN vs. MCI)로 가장 우수했다. M. Richter-Laskowska 외[47]는 소규모 데이터셋에서 MLP(Multi Layer Perceptron)와 SVM(Support Vector Machine), KNN

(K-Nearest Neighbors), RF(Random Forest)등의 여러 머신러닝 기법을 적용한 결과, CN vs. MCI에서는 비선형 SVM(정확도 69%)이, CN vs. Dementia(이하 D)에서는 RF(정확도 84%)이, CN vs. (MCI + D)에서는 비선형 SVM(정확도 75%)이 최고 성능이었다고 보고했다. A. Alatrany 외[48]는 RF, KNN, Naïve Bayes, SVM으로 전자의무기록 데이터를 이용한 AD 분류 실험을 진행했고 각각 96.4%(AD vs. CN, RF), 78.1%(CN vs. MCI, RF), 78.2%(AD vs. MCI, Naïve Bayes)의 정확도가 나타났음을 발표했다.

2.2 TabTransformer

최근 임상 및 코호트 기반 연구 환경에서는 영상의존도를 낮추고 구조화된 테이블 데이터를 효과적으로 활용하려는 시도가 증가하고 있으며, 이에 따라 테이블형 데이터에 최적화된 딥러닝 모델이 주목받고 있다. 그중 TabTransformer는 범주형 특성을 임베딩 공간으로 사상한 뒤 셀프 어텐션(Self-attention) 구조를 통해 변수 간 비선형 상호작용을 학습하는 아키텍처로, 기존의 트리 기반 모델이 갖는 표현 한계를 보완하면서도 경쟁력 있는 예측 성능을 보고하였다[49]. 이러한 방법론적 발전을 바탕으로, 본 연구에서는 ADNI 임상 테이블 데이터를 대상으로 TabTransformer의 질병 분류 성능을 실증적으로 분석하고 그 적용 타당성을 평가한다.

MCI와 AD는 독립적인 생활 가능 여부 등에서 미세한 차이를 보이며 임상적 특성이 상당히 중첩된다. TabTransformer는 멀티 헤드 셀프 어텐션(Multi-head self-attention)을 통해 각 임상 지표가 다른 변수들과 어떤 맥락에서 결합되는지를 의미론적인 측면에서 학습한다. 예를 들어, 동일한 MMSE 및 CDR 검사 결과이더라도 연령이나 APOE4 유전자 유무에 따라 모델이 부여하는 가중치는 상이할 것이다. 이와 같이 문맥을 고려한 학습은 경계가 모호한 질병 분류에 강점을 보인다. TabTransformer의 구조와 알고리즘에 관한 상세한 설명은 3.2절에서 다룬다.

III. 연구 방법

3.1 사용 데이터

본 연구에서는 ADNI(Alzheimer's Disease Neuroimaging Initiative) 공개 데이터베이스에 포함된 6개의 하위 데이터셋을 활용하였다. ADNI는 표준화된 수집 프로토콜에 따라 임상 정보, 신경영상 자료, 바이오마커 및 유전 정보를 장기간에 걸쳐 구축한 코호트 데이터베이스로, 다양한 분석 연구의 재현성과 신뢰성을 확보할 수 있는 고품질 자료를 제공한다[50]-[52].

표 1. 사용 데이터셋 목록
Table 1. Used data sets list

| CRF name | TBL name | Last update | Features |
|---|-----------|--------------|--------------------------------|
| Diagnostic summary [ADNI1,GO,2,3,4] | DXSUM | Oct 24, 2025 | DIAGNOSIS |
| Clinical dementia rating [ADNI1,GO,2,3,4] | CDR | Oct 24, 2025 | CDRSB* |
| Mini-Mental state examination (MMSE) [ADNI1,GO,2,3,4] | MMSE | Oct 24, 2025 | MMSCORE** |
| Functional assessment questionnaire [ADNI1,GO,2,3,4] | FAQ | Oct 24, 2025 | FAQTOTAL*** |
| ADNIMERGE - key ADNI tables merged into one table [ADNI1,GO,2,3] | ADNIMERGE | Oct 25, 2025 | AGE, APOE4, PTGENDER(GENDER) |
| Vital signs [ADNI1,GO,2,3,4] | VITALS | Oct 24, 2025 | HEIGHT, WEIGHT, VSPULSE(PULSE) |

* Total of CDR test score

** Total of MMSE test score

*** Total of FAQ test score

본 연구에서 활용한 6개의 하위 데이터셋의 구체적인 구성은 표 1에 정리하였다. 서론에서 언급한 주요 임상 평가 도구인 치매척도검사(CDR, Clinical Dementia Rating)와 간이인지기능검사(MMSE), 그리고 아포지단백 E(APOE) 관련 유전 정보를 핵심 변수로 포함하였으며, 환자의 기본 인구학적·신체적 특성(연령, 성별, BMI, 맥박)을 함께 반영하였다. 또 일상생활 수행 능력을 평가하는 FAQ(Functional

Assessment Questionnaire) 지표를 추가적으로 활용하였다. 이를 통해 AD와 관련된 임상적 요인을 다각도로 반영하고, 질병 상태에 영향을 미치는 주요 변수들을 정량적으로 분석하고자 하였다.

여기서 가장 핵심 변수로 사용된 CDRSB, MMSCORE, FAQTOTAL은 실제 임상 현장에서 알츠하이머병의 진단 및 중증도 판정을 결정짓는 지표이다. 각 지표와 진단 판단의 구체적인 관계는 다음과 같다.

CDR은 기억력, 지남력, 판단력 및 문제해결 능력 등 6개의 영역의 기능을 평가하여 합산한 점수이다. 이는 인지 기능 저하뿐 아니라 일상생활수행 능력의 저하 정도를 정량화하며 특히 정상(CN), MCI, 초기 치매를 구분 짓는 임상적 기준으로 활용된다. 즉 치매의 진단 후 그 중증도를 평가하는 성격이 강하다[12][53][54].

MMSE는 전반적인 인지 기능 상태를 신속하게 선별하기 위한 도구이다. 질병의 진행 과정 동안 반복 측정함으로써 시간에 따른 변화를 볼 수 있다. [54][55]

FAQ는 일상생활 수행 능력 저하 정도를 평가하는 지표로, 치매 진단 및 중증도 판단에 보조적으로 활용된다.

각 데이터셋의 정확한 병합을 위해 연구 단계, 피험자 번호, 방문 시점을 기준으로 이루어졌으며, NULL과 같은 결측값은 제거하였다. 한 피험자에 관한 여러 시점의 관찰 값이 중복으로 학습이 되어 과적합을 유도할 수 있어 bl(Baseline) 시점만으로 데이터셋을 구성하였다. 전처리를 완료한 최종 데이터셋은 총 2,108행으로, 데이터셋의 구성은 표 2와 그림 1에서 제시한다.

각 그룹별 환자 수는 MCI 환자군이 957명(45.4%), CN이 770명(36.5%), AD가 381명(18.1%)으로 클래스 불균형 현상이 관찰된다.

인지 기능 및 일상생활 수행 능력을 평가하는 CDR, MMSE, FAQ 지표에서 AD 집단은 다른 그룹과 비교하여 뚜렷한 차이를 나타낸다. 중증도가 높을수록 점수가 증가하는 CDR과 FAQ의 경우 AD군이 가장 높은 평균값을 보였으며, 반대로 점수가 낮을수록 인지 기능 저하가 심각함을 의미하는 MMSE에서는 AD군이 가장 낮은 평균값을 기록하였다.

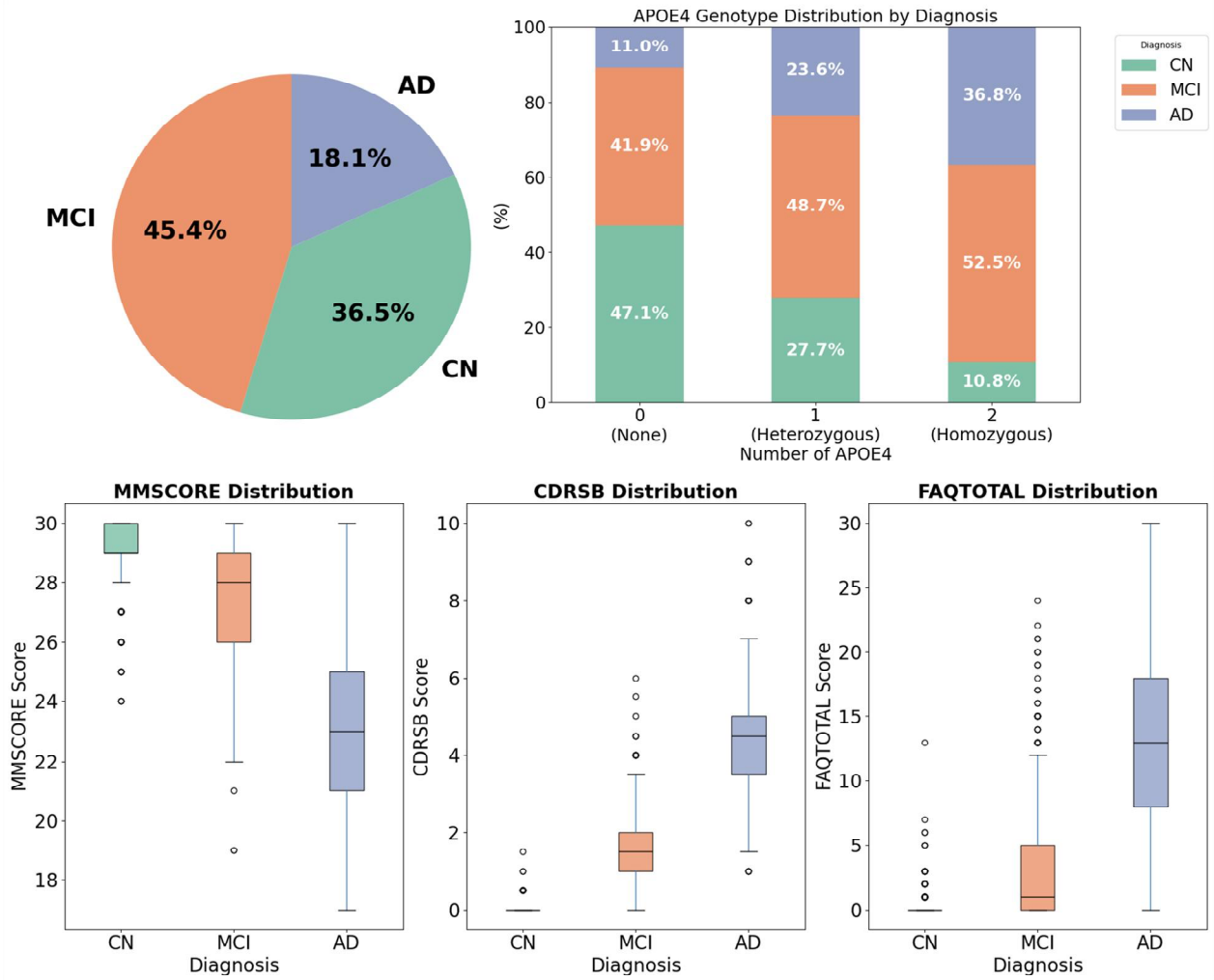


그림 1. 데이터셋 분포
Fig. 1. Distribution of data set

표 2. 데이터셋 구성
Table 2. Distribution of data set

| | CN | MCI | AD |
|---------------------------------|------------------|------------------|------------------|
| N(F/M) | 770 (431/339) | 957 (391/566) | 381 (171/210) |
| AGE (mean(std)) | 73.02 (6.29) | 72.99 (7.53) | 74.86 (7.87) |
| BMI (mean(std)) | 26.92 (4.43) | 26.73 (4.23) | 25.62 (4.11) |
| PULSE (mean(std)) | 65.52 (10.36) | 64.51 (10.16) | 63.85 (9.59) |
| CDRSB (mean(std)) | 0.04 (0.15) | 1.51 (0.90) | 4.40 (1.69) |
| MMSCORE (mean(std)) | 29.10 (1.11) | 27.61 (1.86) | 23.17 (2.18) |
| FAQTOTAL (mean(std)) | 0.19 (0.82) | 3.16 (4.11) | 13.18 (6.89) |

이는 해당 평가 지표들이 질병 단계에 따라 일관된 변별력을 갖고 있음을 시사한다.

또한 AD 집단은 다른 그룹에 비해 상대적으로 큰 분산을 보였는데, 이는 동일한 AD 진단 범주 내에서도 증상의 증중도가 경도에서 중증에 이르기까지 폭넓게 분포하고 있음을 의미한다. 이러한 결과는 AD 환자군 내부의 임상적 이질성이 상당함을 보여주는 근거로 해석될 수 있다.

3.2 모델 설계

본 연구는 테이블 데이터 처리에 특화된 TabTransformer 아키텍처를 기반으로 모델을 설계하였다[56]. 해당 모델은 변수의 관계성을 학습하는 트랜스포머 블록과 MLP로 구성된다.

TabTransformer는 범주형 변수와 수치형 변수를 각각 다른 방식으로 처리한다.

범주형 변수는 컬럼 임베딩(Column embedding)을 하여 각 컬럼의 고유한 값에 컬럼 식별자까지 더해진 문맥화된 벡터를 생성한다. 데이터에 m 개의 범주형 변수 $x_i (i \in 1, 2, 3, \dots, m)$ 가 존재할 때, 각 변수가 d_i 개의 클래스를 가진다면, 결측치 클래스를 포함하여 총 $d_i + 1$ 개의 임베딩 벡터를 생성한다. 여기서 x_i 의 임베딩 참조 테이블을 e_{ϕ_i} 이라 하면, e_{ϕ_i} 는 식 (1)과 같이 컬럼 식별자 c_{ϕ_i} 와 값 벡터 $w_{\phi_{ij}}$ 의 단순 연결로 정의된다.

$$e_{\phi_i}(j) = [c_{\phi_i}, w_{\phi_{ij}}] (x_i = j \in 0, 1, 2, \dots, d_i) \quad (1)$$

생성된 컬럼 임베딩의 집합 $E_{\phi}(x_{cat})$ 은 트랜스포머 블록 f_{θ} 을 순차적으로 N 번 통과하고, 각 임베딩은 다른 컬럼의 임베딩을 반복적으로 참조하여 식 (2)와 같은 문맥화된 임베딩(Contextual Embedding) h_i 으로 변환된다.

$$f_{\theta}(e_{\phi_1}(x_1), e_{\phi_2}(x_2), \dots, e_{\phi_m}(x_m)) = h_1, h_2, \dots, h_m \quad (2)$$

수치형 변수는 레이어 정규화(Layer normalization)를 거친 후, 앞서 생성된 범주형 변수의 문맥화된 임베딩과 단순 결합되어 MLP 층인 g_{ψ} 에 입력된다. 최종적으로 모델은 식 (3)의 손실 함수 \mathcal{L} 로 최적화된다. 앞서 설명한 TabTransformer의 구조를 그림 2에서 간략화하여 제시한다.

$$\mathcal{L}(x, y) = H(g_{\psi}(f_{\theta}(E_{\phi}(x_{cat})), x_{cont}), y) \quad (3)$$

위와 같은 구조를 바탕으로 다음과 같이 모델을 설계했다. 입력 테이블의 형태는 결측치를 제거하고 각 환자의 환자 번호, 방문시기를 기준으로 통합한 데이터다. 독립 변수는 CDRSB, MMSCORE, FAQTOTAL, APOE4, 나이(AGE), 성별(GENDER), BMI, 맥박(PULSE)으로 총 8개이며, APOE4 유전자 개수와 성별은 범주형으로 간주하였다. 종속 변수는 피험자의 질병(DIAGNOSIS)이다.

본 연구에서 설계한 TabTransformer의 전체 모델 구조 및 실험 파이프라인은 그림 3에 도식화하였다. 각 범주형 변수의 임베딩 차원은 32로 설정하였다. 모델의 깊이는 2개의 멀티헤드 어텐션과 2개의 트랜스포머 블록으로 구성하였으며, 분류기를 위한 MLP는 2개 층으로 설계하였다. 과적합 방지를 위해 드롭아웃 비율은 0.1로 설정하였다. 또한 데이터 불균형 문제를 완화하기 위해 클래스 가중치(Class weight)를 손실 함수에 반영하였으며, 이는 오분류 클래스에 대해 더 큰 손실 값을 부여하는 방식으로 구현하였다[57][58].

모델의 일반화 성능을 검증하기 위해 층화 10겹 교차 검증(Stratified 10-fold cross validation)을 적용하였다. 이는 각 폴드 내 클래스 비율을 원본 데이터 분포와 유사하게 유지함으로써 평가의 객관성을 확보하는 방식이다. 더불어, 각 폴드 내 독립적으로 인코딩과 클래스 가중치, 정규화, 학습 파라미터가 적용되도록 설계하여 모델의 신뢰성과 예측 성능을 정확히 판별할 수 있도록 한다. 이에 따른 학습용 데이터셋, 검증용 데이터셋, 테스트 데이터셋의 비율은 다음과 같다. 10%를 테스트용으로 고정하고 9 폴드를 학습용 및 검증용으로 분리한 후, 이 중 10%를 검증용으로 사용한다.

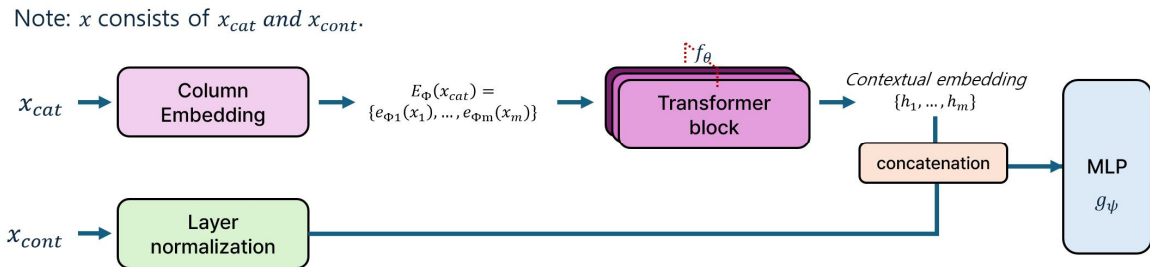


그림 2. TabTransformer 구조
Fig. 2. TabTransformer architecture

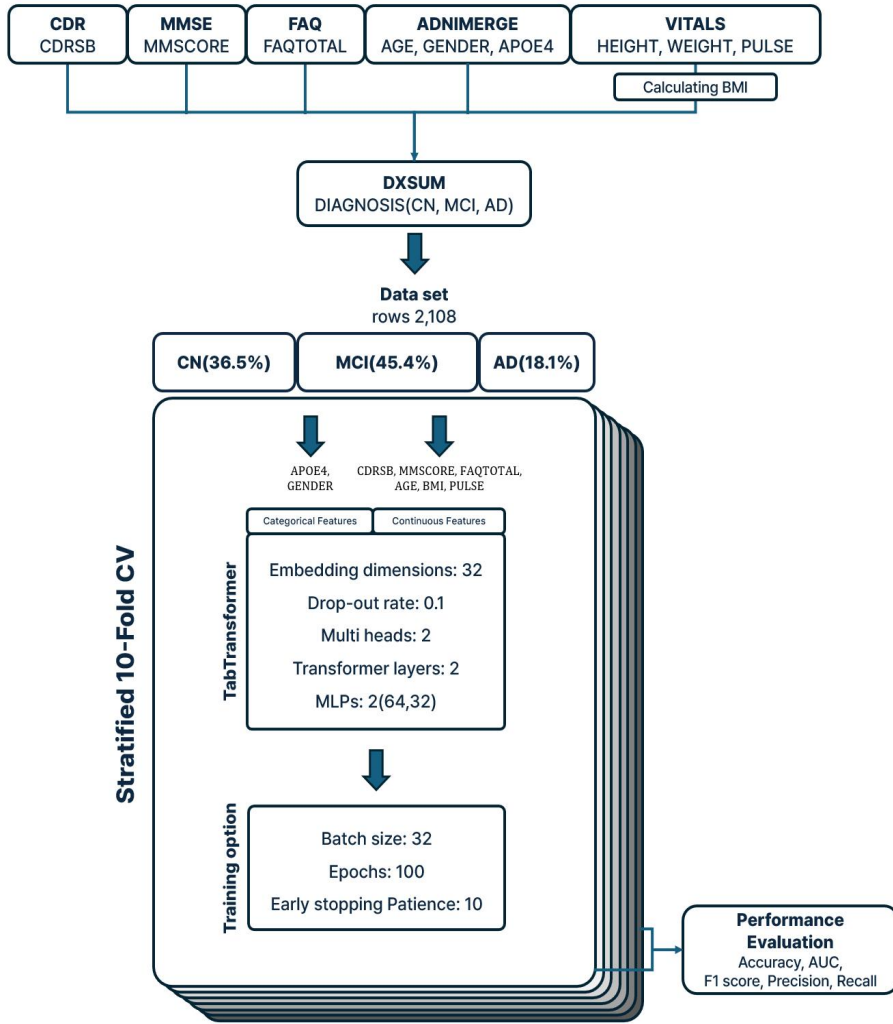


그림 3. 모델 설계 구조도
Fig. 3. Architecture diagram

학습 과정에서 배치 크기는 32, 최대 학습 에폭은 100으로 설정하였고 조기 종료(Early stopping) 기법을 적용했다. 구체적으로 매 에폭마다 검증 손실을 모니터링하였으며, 해당 지표가 10회 연속으로 개선되지 않을 경우 학습을 즉시 중단하도록 하였다. 이때 조기 종료 판단을 위해 검증 데이터를 활용하여 정보 누수를 방지하였다.

최종 성능 평가는 정확도, 정밀도, 재현율, F1 Score, AUC의 평균값을 기준으로 산출하였으며, 각 지표의 정의는 식 (4)~(7)에 제시하였다.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

IV. 실험 결과

4.1 성능 평가

제안된 모델의 분류 성능을 체계적으로 검증하기 위하여 다양한 임상적 난이도를 반영한 실험 환경을 구성하였다. 평가 대상은 (a) CN vs. MCI, (b) MCI vs. AD, (c) CN vs. AD의 세 가지 조합별 이진 분류 과제이다. 각 과제는 임상적 구분 난이도가 상이하며, 이를 단일 다중 분류 문제로 통합할 경우

상대적으로 구분이 용이한 집단의 영향으로 전체 성능이 과대평가될 가능성이 있다. 일반적으로 CN vs. AD는 임상적 차이가 비교적 명확한 반면, MCI vs. AD는 인지 기능 및 생활 능력 지표가 중첩되어 분류 난이도가 높은 과제로 알려져 있다. 이에 각 집단 조합에 대한 모델의 판별 능력을 독립적으로 평가하기 위하여 이진 분류 실험을 설계하였다.

또한 어텐션 메커니즘의 기여도를 검증하기 위해 전통적인 머신러닝(ML) 모델 및 기본 딥러닝 모델인 다층 퍼셉트론(MLP)과의 성능 비교를 수행하였다. 비교 모델의 최적 성능을 도출하기 위해 그리드 서치(Grid search)를 통한 하이퍼파라미터 튜닝을 진행하였으며, 최종 채택된 하이퍼파라미터는 다음 표 3과 같다.

비교 모델들 역시 일반화 성능 확보와 공정한 대조 환경 구축을 위해 층화 10겹 교차 검증과 조기 종료 로직을 동일하게 적용하였다. 특히 정보 누수(Data leakage)를 방지하기 위해 정규화, 인코딩, 클래스 가중치 적용, 하이퍼파라미터 최적화 등 모든 전처리 과정은 전체 데이터셋이 아닌 각 폴드(Fold)의 학습 데이터만을 기준으로 독립적으로 수행되었다. 이는 검증 데이터의 통계 정보가 학습 과정에 사전에 반영되는 편향을 엄격히 차단함으로써, 모델

평가의 객관성과 재현성을 확보하기 위함이다. 여기서 MLP는 TabTransformer에서 트랜스포머 블록의 효과를 확인하기 위한 대조 실험의 대조군이다. 모든 비교 실험에서는 클래스 가중치 적용 및 층화 10겹 교차 검증을 동일하게 수행하여 실험 조건을 일관되게 유지함으로써 평가의 공정성과 재현성을 확보하였다.

표 3. 머신러닝 모델 및 MLP의 하이퍼파라미터
Table 3. Hyperparameter in ML and MLP

| Models | Hyper parameters | Values |
|---------------|-----------------------|------------------|
| Random Forest | n_estimators | 1000 |
| | max_depth | 10 |
| | min_samples_split | 2 |
| | min_samples_leaf | 2 |
| | class_weight | balanced |
| XGBoost | n_estimators | 1000 |
| | max_depth | 8 |
| | learning_rate | 0.1 |
| | min_child_weight | 6 |
| | scale_pos_weight | ratio |
| | early_stopping_rounds | 10 |
| MLP | embedding_dim | 16 |
| | dropout_rate | 0.1 |
| | hidden layers | 2(32, 32) |
| | activation | relu and sigmoid |

표 4. 모델의 성능(정확도, AUC, F1 점수, 재현율, 정밀도)
Table 4. Performances of models(Accuracy, AUC, F1 Score, Recall and Precision)

| | | (a) CN vs. MCI | (b) MCI vs. AD | (c) CN vs. AD |
|-----------|----------------|------------------------|------------------------|------------------------|
| Accuracy | TabTransformer | 0.9618 ± 0.0208 | 0.9275 ± 0.0194 | 0.9983 ± 0.0034 |
| | MLP | 0.9589 ± 0.0135 | 0.9200 ± 0.0171 | 0.9965 ± 0.0080 |
| | Random Forest | 0.9606 ± 0.0096 | 0.9275 ± 0.0199 | 0.9983 ± 0.0035 |
| | XGBoost | 0.9491 ± 0.0133 | 0.9208 ± 0.0336 | 0.9974 ± 0.0040 |
| AUC | TabTransformer | 0.9933 ± 0.0044 | 0.9743 ± 0.0136 | 1.0000 ± 0.0001 |
| | MLP | 0.9932 ± 0.0026 | 0.9743 ± 0.0077 | 0.9999 ± 0.0003 |
| | Random Forest | 0.9911 ± 0.0034 | 0.9729 ± 0.0127 | 1.0000 ± 0.0000 |
| | XGBoost | 0.9888 ± 0.0057 | 0.9693 ± 0.0181 | 1.0000 ± 0.0000 |
| F1 Score | TabTransformer | 0.9654 ± 0.0195 | 0.8771 ± 0.0315 | 0.9974 ± 0.0052 |
| | MLP | 0.9630 ± 0.0125 | 0.8670 ± 0.0222 | 0.9947 ± 0.0122 |
| | Random Forest | 0.9654 ± 0.0080 | 0.8736 ± 0.0351 | 0.9974 ± 0.0052 |
| | XGBoost | 0.9542 ± 0.0121 | 0.8628 ± 0.0578 | 0.9961 ± 0.0060 |
| Recall | TabTransformer | 0.9674 ± 0.0453 | 0.9042 ± 0.0845 | 0.9974 ± 0.0129 |
| | MLP | 0.9662 ± 0.0283 | 0.9035 ± 0.0824 | 0.9921 ± 0.0256 |
| | Random Forest | 0.9875 ± 0.0112 | 0.8818 ± 0.0504 | 1.0000 ± 0.0000 |
| | XGBoost | 0.9593 ± 0.0262 | 0.8739 ± 0.0633 | 1.0000 ± 0.0000 |
| Precision | TabTransformer | 0.9634 ± 0.0226 | 0.8516 ± 0.0458 | 0.9974 ± 0.0076 |
| | MLP | 0.9598 ± 0.0127 | 0.8333 ± 0.0568 | 0.9973 ± 0.0081 |
| | Random Forest | 0.9446 ± 0.0193 | 0.8680 ± 0.0455 | 0.9949 ± 0.0103 |
| | XGBoost | 0.9500 ± 0.0189 | 0.8541 ± 0.0660 | 0.9923 ± 0.0118 |

표 4에 제시된 결과에 따르면, 전반적으로 TabTransformer는 모든 분류 설정에서 비교 모델 대비 우수한 성능을 나타냈다. 특히 어블레이션 실험에서 대조군으로 설정한 MLP 모델보다 주요 평가 지표 전반에서 높은 값을 기록하였다. 표 4(c)에 해당하는 CN vs. AD 분류 설정에서는 정확도 0.9983, AUC가 1에 근접하는 수준으로 산출되었으며, 이는 두 집단 간 임상적 차이가 비교적 명확하여 대부분의 모델이 높은 분류 성능을 보인 결과로 해석된다.

한편, 분류 난이도가 가장 높은 것으로 알려진 표 4(b) 설정(MCI vs. AD)에서는 TabTransformer가 정확도, AUC, F1 점수 모두에서 최고 성능을 기록하였다. 이는 임상적 특성이 중첩되는 집단 간 구분 과정에서 변수 간 복합적 상호작용을 효과적으로

학습함으로써, 제안 모델이 보다 정교한 판별 능력을 확보하였음을 시사한다.

각 모델의 분류 성능을 ROC-AUC 곡선으로 시각화한 결과는 그림 4와 같다. 그림 4는 세 가지 이진 분류 과제(CN vs. MCI, MCI vs. AD, CN vs. AD)에 대한 TabTransformer, MLP, Random Forest, XGBoost의 ROC 곡선과 평균 AUC를 함께 제시한다. CN vs. AD 과제에서는 모든 모델의 곡선이 좌상단에 근접하여 AUC가 1에 가까운 우수한 분류 성능을 보인 반면, 분류 난이도가 가장 높은 MCI vs. AD 과제에서는 모델 간 곡선의 차이가 상대적으로 뚜렷하게 나타났다. 특히 해당 과제에서 TabTransformer가 가장 높은 평균 AUC(0.9743)를 기록하여, 임상적 특성이 중첩되는 집단에 대한 판별력에서 비교 모델 대비 우위를 보였다.

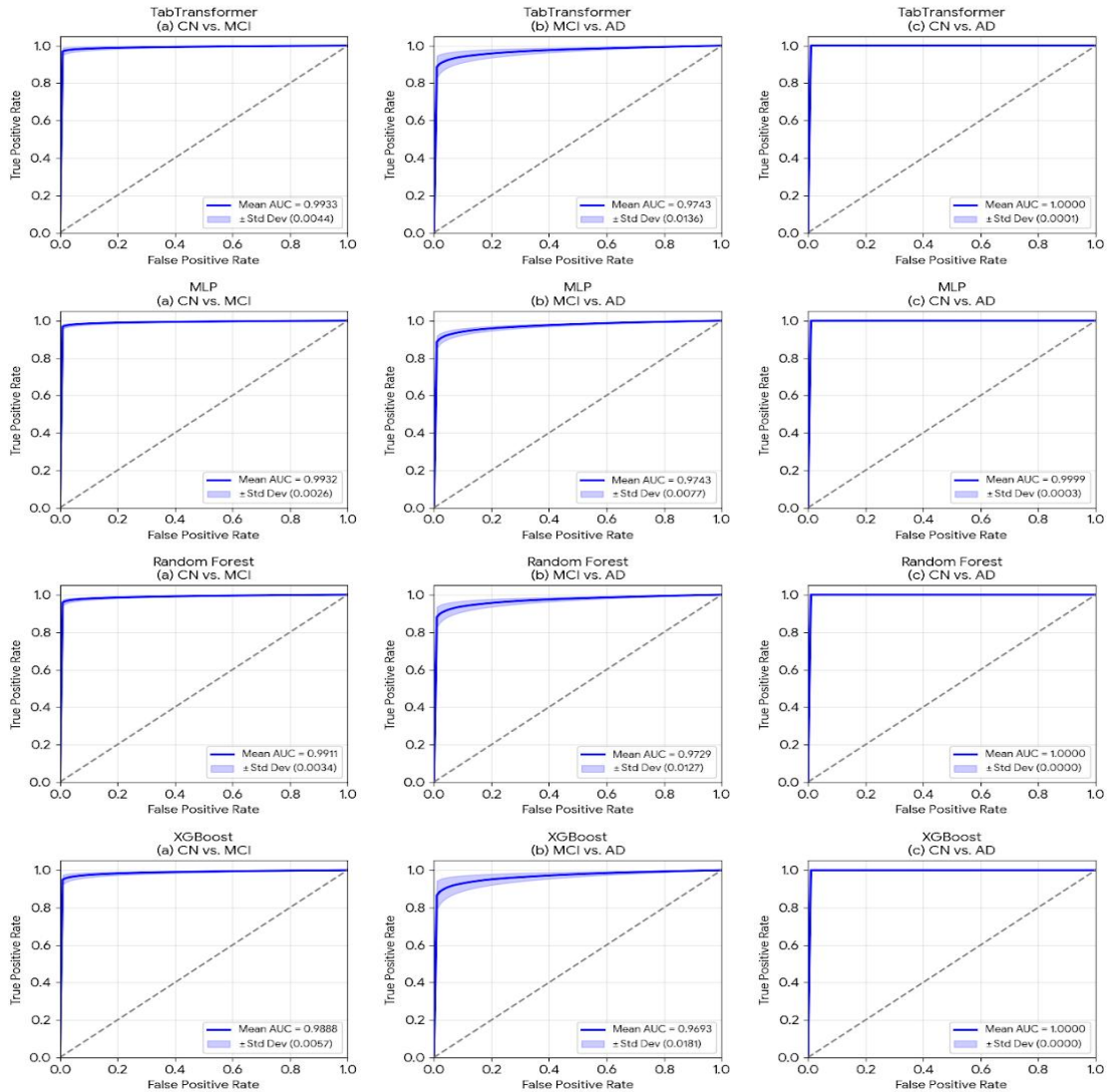


그림 4. ROC-AUC 곡선
Fig. 4. ROC-AUC Curves

추가로 본 연구는 진단 지표가 모델 성능에 미치는 지배적인 영향을 확인하고 정보 누수 우려를 검토하기 위해, 해당 변수들을 제외한 추가 실험을 수행하였다. AUC의 감소도를 기준으로 알츠하이머병과 관련된 직접적 연관 변수(CDRSB, MMSCORE, FAQTOTAL)를 하나씩 제외하거나 모두 제거하는 보조 실험을 수행했다. 그 결과, 절대적인 성능은 표 5와 같이 하락하였다. 표 5(a)는 CDRSB 제거 시 가장 많이 하락하였고, 표 5(b)는 각 세 개의 변수의 유무에 따라 0.86대로 하락하는 경향을 보였다. 표 5(c)는 세 개의 변수의 유무에 따라 큰 변화 없이 미미한 차이를 보였다. 세 변수를 모두 제거 시에는 표 5(b)의 경우가 가장 많이 하락했으므로 보아 AD 관련 변수에 가장 의존하는 단계 분류 과제임을 짐작할 수 있다.

종합적으로 제안 모델의 성능에 관하여 다음과 같이 논의할 수 있다. 본 연구에서 사용한 CDRSB, MMSCORE, FAQTOTAL 등은 실제 진단을 내릴 때 사용하는 핵심 기준으로써, 이 변수들이 진단 결과와 매우 밀접한 연관성을 가지고 있기에 제안 모델이 강력한 변별력을 가졌다고 할 수 있다. 또한 제안 모델의 어텐션 메커니즘은 단순 트리 기반 모델과는 달리 변수 간의 복합적인 상호작용을 학습하므로 성능 향상의 주된 요인이라 볼 수 있다.

4.2 선행 연구와의 비교

본 연구는 AD 단계 분류를 수행한 선행 연구와의 비교를 통해 제안 모델의 유효성을 검증하였으며, 그 결과를 표 6에 종합하였다. F. Abuhantash et al.[46]는 원본 피쳐 집합(Original feature sets)을 활용한 경우와 인구통계학적 변수, 인지 검사 지표, 동반질환 정보 등 서로 다른 특성 조합을 적용한

어블레이션 실험을 수행하여 입력 특성 구성에 따라 분류 성능이 유의하게 변화함을 보고하였다. 세 가지 이진 분류 설정에서 보고된 최대 성능은 (a) 정확도 0.94, AUC 0.99, F1 점수 0.94, (b) 정확도 0.93, AUC 0.98, (c) 정확도 0.99, AUC 0.99, F1 점수 0.99이다. 이에 비해 본 연구는 (b) 설정에서 정밀도 측면에서는 일부 낮은 값을 보였으나, AUC 0.97을 기록하여 판별 능력 측면에서는 경쟁력 있는 성능을 확보하였다.

표 6. 선행 연구와의 비교
Table 6. Comparison of prior studies

| | | (a) | (b) | (c) |
|------------------------------|-----|-------------|-------------|-------------|
| F. Abuhan-tash et al.[46] | ACC | 0.94 | 0.93 | 0.99 |
| | AUC | 0.99 | 0.98 | 0.99 |
| | F1 | 0.94 | 0.93 | 0.99 |
| M. Priyadh-rshini et al.[59] | ACC | None | None | 0.95 |
| | AUC | None | None | 0.98 |
| | F1 | None | None | 0.96 |
| S. Bit et al. [60] | ACC | 0.75 | 0.81 | 0.93 |
| | AUC | None | None | None |
| | F1 | 0.80 | 0.87 | 0.96 |
| S. Wang et al.[61] | ACC | 0.83 | 0.93 | 0.97 |
| | AUC | None | None | None |
| | F1 | 0.82 | 0.94 | 0.97 |
| This study | ACC | 0.96 | 0.93 | 1.00 |
| | AUC | 0.99 | 0.97 | 1.00 |
| | F1 | 0.97 | 0.88 | 1.00 |

표 5. TabTransformer에서 주요 변수를 제거한 보조 실험
Table 5. Result of Feature ablation study in TabTransformer

| | | (a) CN vs. MCI | (b) MCI vs. AD | (c) CN vs. AD |
|-----|----------|-----------------|-----------------|-----------------|
| AUC | Original | 0.9933 ± 0.0044 | 0.9743 ± 0.0136 | 1.0000 ± 0.0001 |
| | CDRSB | 0.8664 ± 0.0349 | 0.8688 ± 0.0207 | 0.9996 ± 0.0006 |
| | MMSCORE | 0.9919 ± 0.0054 | 0.8688 ± 0.0217 | 0.9994 ± 0.0009 |
| | FAQTOTAL | 0.9904 ± 0.0062 | 0.8659 ± 0.0241 | 0.9994 ± 0.0010 |
| | ALL | 0.6230 ± 0.0339 | 0.6042 ± 0.0449 | 0.7536 ± 0.0337 |

M. Priyadharshini et al.[59]는 테이블 데이터를 기반으로 Enhanced TabTransformer 모델을 적용하여 AD와 CN을 구분하였다. 해당 연구가 클래스 불균형 완화를 위해 SMOTE(Synthetic Minority Over-sampling Technique)를 적용한 것과 달리, 본 연구는 원본 데이터의 분포를 유지한 상태에서 학습을 수행하였으며, 그럼에도 불구하고 향상된 분류 성능을 달성하였다.

S. Bit et al.[60]는 ADNI와 OASIS(Open Access Series of Imaging Studies) 데이터셋의 3D sMRI 영상만을 입력으로 사용하여 VAE(Variational Autoencoder) 기반 잠재 특징을 추출한 후 LGBM(Light Gradient Boosting Machine) 등 머신러닝 분류기를 적용하였다. 보고된 최적 성능은 (a) 정확도 0.75, F1 점수 0.80, (b) 정확도 0.81, F1 점수 0.87, (c) 정확도 0.93, F1 점수 0.96으로, 전반적으로 본 연구 대비 낮은 수준을 보였다. 이는 영상 기반 단일 모달 특징보다 구조화된 임상 테이블 데이터가 해당 분류 문제에서 높은 변별력을 가질 가능성을 시사한다.

한편, S. Wang et al.[61]는 종단적 MRI 및 PET 영상, 인지 기능 검사, 유전자 검사 등을 결합한 멀티모달 데이터를 활용하여 hypergraph 정규화와 multi-task feature selection을 기반으로 multi-kernel SVM을 적용하였다. 본 연구는 (b) 설정에서 AUC 0.97을 기록하여 판별 능력 측면에서는 경쟁력 있는 결과를 보였으나, F1 점수에서는 멀티모달 접근 대비 다소 낮은 수치를 나타냈다. 종합적으로 볼 때, 본 연구는 단일 모달 영상 기반 접근에 비해서는 우수한 성능을 확보하였으나, 멀티모달 정보의 통합이 이루어지지 않았다는 점에서 정밀도 측면의 한계가 존재함을 확인하였다.

4.3 중요도 분석

TabTransformer를 포함한 대부분의 딥러닝 모델은 내부 의사결정 과정이 명시적으로 드러나지 않는 블랙박스(Black-box) 특성을 가지므로, 모델 해석을 위해 별도의 사후적 분석 기법이 요구된다. 이에 본 연구에서는 변수의 상대적 기여도를 평가하기 위하여 순열 중요도(Permutation importance) 기법을 적용하였다. 순열 중요도는 특정 변수를 무작위로 재배열한 뒤 모델 성능의 변화를 측정하고, 이를 원래 성능과 비교함으로써 해당 변수가 예측에 기여하는 정도를 정량화하는 방식이다[62][63]. 성능 저하 폭이 클수록 모델이 해당 변수를 핵심적으로 활용하고 있음을 의미한다. 이 방법은 모델 내부 구조를 직접 해석하지 않고 입력 데이터를 변형하여 평가

를 수행하므로, 모델 구조에 독립적인 해석 기법이라는 장점을 가진다.

본 연구에서 순열 중요도는 층화 10겹 교차 검증 결과의 평균 정확도 감소량을 기준으로 산출하였다. 표 7과 그림 5는 각 이진 단계 분류 모델에 대한 상위 5개 변수의 순열 중요도를 제시한다.

표 7(a)에서는 CDR의 정확도 감소량이 약 42%로 산출되어 다른 변수 대비 현저히 높은 기여도를 나타냈다. 이는 CN과 MCI를 구분하는 과정에서 CDR 지표가 핵심적인 판별 변수로 작용하고 있음을 의미한다. 또한 표 7(c)에서도 CDR이 가장 높은 정확도 감소량을 기록하여, 임상적 차이가 비교적 명확한 집단 간 분류에서 CDR의 영향력이 지배적임을 확인할 수 있다. 반면, 분류 난이도가 높은 표 7(b)에서는 MMSE(약 12%)와 CDR(약 9.7%)이 상대적으로 높은 중요도를 보였으나, 특정 단일 변수가 압도적인 영향력을 나타내지는 않았다. 이는 임상적 경계가 중첩되는 MCI와 AD 구분에서 복수의 임상 지표가 상호보완적으로 작용하고 있음을 시사한다.

표 7. 순열 중요도

Table 7. Permutation importance

| (a) CN vs. MCI | | |
|------------------------------------|----------|------------|
| Baseline accuracy: 0.9618 ± 0.0208 | | |
| No. | Feature | Importance |
| 1 | CDRSB | 0.4221 |
| 2 | GENDER | 0.0093 |
| 3 | FAQTOTAL | 0.0081 |
| 4 | MMSCORE | 0.0075 |
| 5 | BMI | 0.0035 |
| (b) MCI vs. AD | | |
| Baseline accuracy: 0.9275 ± 0.0194 | | |
| No. | Feature | Importance |
| 1 | MMSCORE | 0.1203 |
| 2 | CDRSB | 0.0972 |
| 3 | FAQTOTAL | 0.0157 |
| 4 | GENDER | 0.0060 |
| 5 | PULSE | 0.0015 |
| (c) CN vs. AD | | |
| Baseline accuracy: 0.9983 ± 0.0034 | | |
| No. | Feature | Importance |
| 1 | CDRSB | 0.1677 |
| 2 | MMSCORE | 0.0868 |
| 3 | FAQTOTAL | 0.0790 |
| 4 | APOE4 | 0.0017 |
| 5 | AGE | 0.0008 |

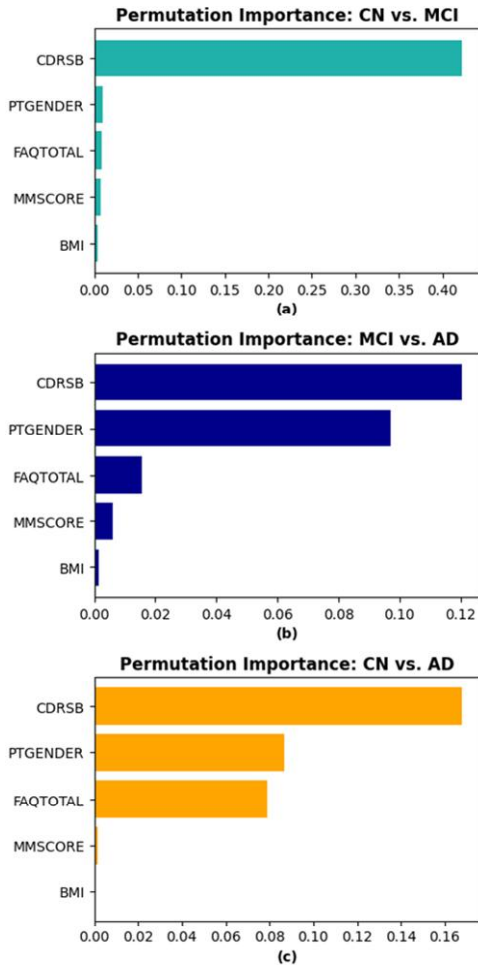


그림 5. 순열 중요도
Fig. 5. Permutation importance

FAQ의 정확도 감소량은 약 1.5%로 비교적 낮은 수준이었으나, 독립적 생활 수행 능력과 관련된 임상적 특성이 보조적 요인으로 반영되었음을 확인할 수 있다. 전반적으로 임상 평가 변수들이 주요 예측 요인으로 작용한 반면, 신체적·유전적 변수는 상대적으로 낮은 기여도를 나타냈다.

V. 결론 및 향후 과제

본 연구에서는 ADNI 데이터베이스에 수록된 피험자의 임상 검사 결과와 신체적·유전적 특성 정보를 통합하여 TabTransformer 기반 질병 상태 단계 분류 모델을 구축하였다. 실험 결과, CN vs. MCI에서 96%, MCI vs. AD에서 93%, CN vs. AD에서 99%의 정확도를 기록하였으며, 이는 기존 선행 연

구 대비 향상된 성능에 해당한다. 특히 분류 난이도가 높은 MCI vs. AD 구분에서 기존 머신러닝 모델 대비 우수한 F1 점수와 AUC를 달성함으로써, 어텐션 기반 표현 학습 구조가 테이블 데이터 환경에서 효과적으로 작동함을 실증적으로 확인하였다. 또한 기본 다층 퍼셉트론(MLP) 모델과의 어블레이션 비교를 통해, TabTransformer에 내재된 어텐션 메커니즘이 질병 상태와 입력 변수 간의 복잡적·비선형적 상호작용을 보다 정교하게 포착함을 검증하였다. 이는 단순 피드포워드 기반 모델이 갖는 표현 한계를 보완할 수 있음을 시사하며, 제안 아키텍처의 구조적 타당성을 뒷받침한다.

한편, 순열 중요도 분석 결과, 세 가지 분류 과제에서 대체적으로 임상 검사와 관련된 변수가 유의미한 영향을 미치는 것으로 확인되었다. CN vs. MCI 과제에서는 CDRSB의 성능 감소량이 약 42.2%로 압도적인 기여도를 보였으며, MCI vs. AD 과제에서는 MMSE(약 12%)와 CDRSB(약 9.7%)가, CN vs. AD 과제에서는 CDRSB(약 16.8%)가 핵심 지표로 작용하였다. 이 결과는 실제 현장의 진단 기준이 모델에도 적절히 반영되었음을 나타내므로 제안된 모델의 신뢰성을 뒷받침하는 근거가 된다.

본 연구는 데이터 편향을 최소화하고 무결성을 확보한 코호트 자료를 기반으로 분석을 수행하였다는 점에서 방법론적 신뢰성을 확보하였으며, 변수 기여도 분석을 통해 모델의 해석 가능성을 함께 제시하였다는 데 의의가 있다. 그러나 본 연구는 특정 시점에서 수집된 단면(Cross-sectional) 데이터를 활용하여 실험을 수행하였으므로, 개별 환자의 질병 진행 경로나 잠재적 발병 가능성을 동적으로 추정하는 데에는 한계가 존재한다. 향후 연구에서는 종단적(Longitudinal) 코호트 데이터를 활용하여 질병의 전이 과정 및 발병 위험도를 학습하는 조기 예측 모델로 확장하고자 한다. 또한 임상적으로 증상 및 인지 특성이 중첩되는 MCI와 AD 간 구분 성능을 보다 정교화함으로써, 모델의 판별력을 강화하는 방향으로 개선을 추진할 계획이다.

또한 본 연구 결과는 사용 데이터셋과 입력 특징 조합에서 최적화된 결과일 가능성이 있으며, 또 다른 코호트나 외부 기관 자료에 대해서는 동일한 수

준의 성능이 나타나지 않을 수 있음을 명시한다. 추후 독립적인 외부 코호트, 다기관 데이터셋을 이용한 검증을 통해 더욱 정교한 모델을 제안할 것이다.

References

- [1] L. Rizzi, I. Rosset, and M. Roriz-Cruz, "Global Epidemiology of Dementia: Alzheimer's and Vascular Types", *BioMed research international*, Vol. 2014, No. 1, Art. no. 908915, Jun. 2014. <https://doi.org/10.1155/2014/908915>.
- [2] Dementia, <https://www.who.int/en/news-room/factsheets/detail/dementia> [accessed: Nov. 30, 2025]
- [3] M. C. Carrillo, L. J. Bain, G. B. Frisoni, and M. W. Weiner "Worldwide Alzheimer's disease neuroimaging initiative", *Alzheimer's & Dementia*, Vol 8, No. 4, pp. 337-342, Jul. 2012. <https://doi.org/10.1016/j.jalz.2012.04.007>.
- [4] Z. S. Khachaturian, "Diagnosis of Alzheimer's Disease", *Arch Neurol*, Vol. 42, No. 11, pp. 1097-1105, Nov. 1985. <https://doi.org/10.1001/archneur.1985.04060100083029>.
- [5] H. Choi and T. Jang, "Effect of Self-management Program on Self-Efficacy, Quality of Life, and Depression in Patients with Mild cognitive impairment", *Journal of Society of Occupational Therapy for the Aged and Dementia*, Vol. 15, No. 1, pp. 23-29, May 2021. <https://doi.org/10.34263/jstod.2021.15.1.23>.
- [6] M. S. Albert, et al., "The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease", *Alzheimer's & Dementia*, Vol. 7, No. 3, pp. 270-279, Apr. 2011. <https://doi.org/10.1016/j.jalz.2011.03.008>.
- [7] R. C. Petersen, "Mild cognitive impairment as a diagnostic entity", *Journal of Internal Medicine*, Vol. 256, No. 3, pp. 183-194, Sep. 2004. <https://doi.org/10.1111/j.1365-2796.2004.01388.x>.
- [8] R. C. Petersen, et al., "Criteria for Mild Cognitive Impairment Due to Alzheimer's Disease in the Community", *Annals of Neurology*, Vol. 74, No. 2, pp. 199-208, Aug. 2013. <https://doi.org/10.1002/ana.23931>.
- [9] H. Kim, et al., "Effect of 12-week exercise program on β -amyloid, BDNF of the elderly with mild cognitive impairment", *Journal of Converging Sport and Exercise Sciences*, Vol. 18, No. 1, pp. 19-26, May 2020. <http://doi.org/10.22997/jcses.2020.18.1.19>.
- [10] B. J. Kelly and R. C. Petersen, "Alzheimer's Disease and Mild Cognitive Impairment", *Neurologic Clinics*, Vol. 25, No. 3, pp. 577-609, Aug. 2007. <https://doi.org/10.1016/j.ncl.2007.03.008>.
- [11] Alzheimer's Association, "2014 Alzheimers disease facts and figures", *Alzheimers Dementia*, Vol. 10, No. 2, pp. 47-92, Mar. 2014. <https://doi.org/10.1016/j.jalz.2014.02.001>.
- [12] Korean Dementia Association, "Clinical Practice Guideline for Dementia (Diagnosis and Evaluation): 2021 Revised Edition", *Dement Neurocogn Discord*, Vol. 21, No. 1, pp. 42-44, Jan. 2022. <https://doi.org/10.12779/dnd.2022.21.1.42>.
- [13] S. Park, et al., "Diagnostic Tools for Alzheimer's Disease: A Narrative Review Based on Our Own Research Experience", *Dement Neurocogn Discord*, Vol. 22, No. 1, pp. 16-27, Feb. 2023. <https://doi.org/10.12779/dnd.2023.22.1.16>.
- [14] F. Riahi, et al., "Comparison of PET/CT and PET/MRI in central nervous system tumors, a narrative review", *Int J Physiol Pathophysiol Pharmacol*, Vol. 16, No. 4, pp. 89-95, Aug. 2024. <https://doi.org/10.62347/UMDS1961>.
- [15] Y. Yim and W. Moon, "Diagnosis and Treatment of Alzheimer's Disease: Current Update", *Journal of the Korean Society of Radiology*. Vol. 86, No. 1, pp. 6-16, Jan. 2025. <https://doi.org/10.3348/jksr.2024.0146>.

- [16] E. H. Corder, et al., "Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families", *Science*, Vol. 261. No. 5123, pp. 921-923, Aug. 1993. <https://doi.org/10.1126/science.8346443>.
- [17] J. Fortea, et al., "APOE4 homozygosity represents a distinct genetic form of Alzheimer's disease", *Nature Medicine*, Vol. 30, No. 5, pp. 1284-1291, May 2024. <https://doi.org/10.1038/s41591-024-02931-w>.
- [18] M. A. Husain, et al., "APOE and Alzheimer's Disease: From Lipid Transport to Physiopathology and Therapeutics", *Frontiers in Neuroscience*, Vol. 15, Art. no. 630502, Feb. 2021. <https://doi.org/10.3389/fnins.2021.630502>.
- [19] B. R. Troutwine, et al., "Apolipoprotein E and Alzheimer's disease", *Acta Pharmaceutica Sinica B*, Vol. 12, No. 2, pp. 496-510, Oct. 2021. <https://doi.org/10.1016/j.apsb.2021.10.002>.
- [20] M. E. Belloy, V. Napolioni, and M. D. Greicius, "A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward", *Neuron*, Vol. 101, No. 5, pp. 820-838, Mar. 2019. <https://doi.org/10.1016/j.neuron.2019.01.056>.
- [21] M. Pires and A. C. Rego, "ApoE4 and Alzheimer's Disease Pathogenesis—Mitochondrial Deregulation and Targeted Therapeutic Strategies", *International Journal of Molecular Sciences*, Vol. 24, No. 1, Art. no. 778, Jan. 2023. <https://doi.org/10.3390/ijms24010778>.
- [22] G. B. Frisoni, et al., "The clinical use of structural MRI in Alzheimer disease", *Nat Rev Neurol*, Vol. 6, No. 2, pp. 67-77, Feb. 2010. <https://doi.org/10.1038/nrneurol.2009.215>.
- [23] E. JR. Beek, et al., "Value of MRI in Medicine: More Than Just Another Test?", *Journal of magnetic resonance imaging*, Vol. 49, No. 7, pp. e14-e25, Jun. 2019. <https://doi.org/10.1002/jmri.26211>.
- [24] Y. Liu, et al., "A low-cost and shielding-free ultra-low-field brain MRI scanner", *Nat Commun*, Vol. 12, No. 1, Art. no. 7238, Dec. 2021. <https://doi.org/10.1038/s41467-021-27317-1>.
- [25] S. Y. Huang, et al., "Portable Low-Cost MRI System Based on Permanent Magnets/Magnet Arrays", *Investigative Magnetic Resonance Imaging*, Vol. 23, No. 3, pp. 179-201, Sep. 2023. <https://doi.org/10.13104/imri.2019.23.3.179>.
- [26] S. Murali, et al., "Bringing MRI to low- and middle-income countries: Directions, challenges and potential solutions", *NMR in Biomedicine*, Vol. 37, No. 7, Art. no. 4992, Jul. 2024. <https://doi.org/10.1002/nbm.4992>.
- [27] J. Ha and K. Kim, "Neighborhood-Regularized Matrix Factorization for lncRNA-Disease Association Identification", *International Journal of Molecular Sciences*, Vol. 26, No. 9, Art. no. 4283, Jan. 2025. <https://doi.org/10.3390/ijms26094283>.
- [28] K. Kim and J. Ha, "GMFLDA: improved prediction of lncRNA-disease association via graph convolutional network", *IEEE Access*, Vol. 13, pp. 85330-85341, May 2025. <https://doi.org/10.1109/ACCESS.2025.3568461>.
- [29] J. Ha, "DeepWalk-based graph embeddings for miRNA-disease association prediction using deep neural network", *Biomedicines*, Vol. 13, No. 3, Art. no. 536, Feb. 2025. <https://doi.org/10.3390/biomedicines13030536>.
- [30] J. Ha, "Graph Convolutional Network with Neural Collaborative Filtering for Predicting miRNA-Disease Association", *Biomedicines*, Vol. 13, No. 1, Art. no. 136, Jan. 2025. <https://doi.org/10.3390/biomedicines13010136>.
- [31] J. Ha, "MDMF: Predicting miRNA-Disease Association Based on Matrix Factorization with Disease Similarity Constraint", *Journal of Personalized Medicine*, Vol. 12, No. 6, Art. no. 885, Jun. 2022. <https://doi.org/10.3390/jpm12060885>.
- [32] J. Ha and S. Park, "NCMD: Node2vec-Based Neural Collaborative Filtering for Predicting MiRNA-Disease Association", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 20, No. 2, pp. 1257-1268, Mar.

2023. <https://doi.org/10.1109/TCBB.2022.3191972>.
- [33] J. Ha, "SMAP: Similarity-based matrix factorization framework for inferring miRNA-disease association", *Knowledge-Based Systems*, Vol. 263, Art. no. 110295, Mar. 2023. <https://doi.org/10.1016/j.knosys.2023.110295>.
- [34] J. Ha, "LncRNA Expression Profile-Based Matrix Factorization for Predicting lncRNA- Disease Association", *IEEE Access*, Vol. 12, pp. 70297-70304, May 2024. <https://doi.org/10.1109/ACCESS.2024.3401005>.
- [35] B. Kim, S. Kim, and J. Ha, "An Ensemble Machine Learning Framework for Identifying LncRNA-Disease Association", *The Journal of Korean Institute of Information Technology*, Vol. 23, No. 6, pp. 41-53, Jun. 2025. <https://doi.org/10.14801/jkiit.2025.23.6.41>.
- [36] S. Kim and J. Ha, "Development of LncRNA-Disease Associations Prediction Model using CNN-RNN Feature Extraction based on LncRNA Sequence Data", *The Journal of Korean Institute of Information Technology*, Vol. 22, No. 10, pp. 1-11, Oct. 2024. <https://doi.org/10.14801/jkiit.2024.22.10.1>.
- [37] J. Ha, "Autoencoder-based Disease-related miRNA Prediction Research using Deep Learning", *The Journal of Korean Institute of Information Technology*, Vol. 20, No. 6, pp. 33-40, Jun. 2022. <https://doi.org/10.14801/jkiit.2022.20.6.33>.
- [38] S. Kim, T. Yun, J. Ryu, and J. Ha, "CNN-GRU and Heterogeneous Graph Convolutional Networks for lncRNA-Disease Association Prediction", *The Journal of Korean Institute of Information Technology*, Vol. 23, No. 11, pp. 1-16, Nov. 2025. <https://doi.org/10.14801/jkiit.2025.23.11.1>.
- [39] T. Yun, S. Kim, J. Ryu, and J. Ha, "Alzheimer's Disease Classification using MRI-based Lightweight CNN-Transformer", *The Journal of Korean Institute of Information Technology*, Vol. 23, No. 11, pp. 181-194, Nov. 2025. <https://doi.org/10.14801/jkiit.2025.23.11.181>.
- [40] J. Ha, "Transfer learning with BioBERT embeddings for lncRNA-disease association prediction", *IEEE Transactions on Computational Biology and Bioinformatics* 2025, Vol. 22, No. 6, pp. 3463-3475. Nov. 2025. <https://doi.org/10.1109/tcbbio.2025.3628675>.
- [41] J. Ha, "SVDTI: Stacked variational autoencoder with SMILES-based drug representations for identifying drug-target interaction", *Neurocomputing* 2026, Vol. 661, Art. no. 131837, Jan. 2026. <https://doi.org/10.1016/j.neucom.2025.131837>.
- [42] D. Yang, R. Wang, and C. Song, "Classification of Alzheimer's Disease Using fMRI-Based Brain Functional Network Data", 2024 6th Asia Symposium on Image Processing(ASIP), Tianjin, China, pp. 97-101. Jun. 2024. <https://doi.org/10.1109/ASIP63198.2024.00025>.
- [43] A. Alorf and M. U. G. Khan, "Multi-label classification of Alzheimer's disease stages from resting-state fMRI-based correlation connectivity data and deep learning", *Computers in Biology and Medicine*, Vol. 151, No. A, pp. 1-16, Nov. 2022. <https://doi.org/10.1016/j.combiomed.2022.106240>.
- [44] J. Noh, J. Kim, and H. Yang, "Classification of Alzheimer's Progression Using fMRI Data", *Sensors*, Vol. 23, No. 14, pp. 6330. Jul. 2023. <https://doi.org/10.3390/s23146330>.
- [45] N. Chauhan and B. Choi, "Classification of Alzheimer's Disease Using Maximal Information Coefficient-Based Functional Connectivity with an Extreme Learning Machine", *Brain Sciences*, Vol. 13, No. 7, Art. no. 1046, Jul. 2023. <https://doi.org/10.3390/brainsci13071046>.
- [46] F. Abuhantash, M. K. A. Hantash, and A. AlShehhi, "Comorbidity-based framework for Alzheimer's disease classification using graph neural networks", *Scientific Reports*, Vol. 14, Art. no. 21061, Sep. 2024. <https://doi.org/10.1038/s41598-024-72321-2>.

- [47] M. Richter-Laskowska, E. Sobotnicka, and A. Bednorz, "Cognitive performance classification of older patients using machine learning and electronic medical records", *Scientific Reports*, Vol. 15, No. 1, Art. no. 6564, Feb. 2025. <https://doi.org/10.1038/s41598-025-90460-y>.
- [48] A. S. Alatrany, W. Khan, A. Hussain, H. Kolivand, and D. Al-Jumeily, "An explainable machine learning approach for Alzheimer's disease classification", *Scientific Reports*, Vol. 14, No. 1, Art. no. 2637, Feb. 2024. <https://doi.org/10.1038/s41598-024-51985-w>.
- [49] X. Huang, A. Khetan, M. Cvitkovic, and Z. S. Karnin, "TabTransformer: Tabular Data Modeling Using Contextual Embeddings", *arXiv preprint arXiv:2012.06678*, pp. 1-7, Dec. 2020. <https://doi.org/10.48550/arXiv.2012.06678>.
- [50] A. W. Toga, et al., "The informatics of ADNI", *Alzheimers Dement*, Vol. 20, No. 10, pp. 7320-7330, Aug. 2024. <https://doi.org/10.1002/alz.14099>.
- [51] M. W. Weiner, et al., "The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception", *Alzheimers Dement*, Vol. 9, No. 5, pp. 111-194, Aug. 2013. <https://doi.org/10.1016/j.jalz.2013.05.1769>.
- [52] R. C. Petersen, et al., "Alzheimer's Disease Neuroimaging Initiative (ADNI)", *Neurology*, Vol. 74, No. 3, pp. 201-209, Jan. 2010. <https://doi.org/10.1212/WNL.0b013e3181eb3e25>.
- [53] S. Choi, et al., "Estimating the validity of the Korean version of expanded Clinical Dementia Rating (CDR) Scale", *Journal of the Korean Neurological Association*, Vol. 19, No. 6, pp. 585-591, Nov. 2001.
- [54] B. Ku, et al., "Clinical practice guideline for dementia by Clinical Research Center for Dementia of South Korea", *Journal of Korean Medical Association*, Vol. 54, No. 8, pp. 861-875, Aug. 2011. <http://doi.org/10.5124/jkma.2011.54.8.861>.
- [55] M. Folstein, et al., "'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician", *Journal of psychiatric research*, Vol. 12, No. 3, pp. 189-198, Nov. 1975. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
- [56] How TabTransformer works, <https://docs.aws.amazon.com/sagemaker/latest/dg/tabtransformer-HowItWorks.html> [accessed: Dec. 27, 2025]
- [57] R. Walsh and M. Tardy, "A Comparison of Techniques for Class Imbalance in Deep Learning Classification of Breast Cancer", *Diagnostics*, Vol. 13, No. 1, Art. no. 67, Dec. 2022. <https://doi.org/10.3390/diagnostics13010067>.
- [58] K. R. M. Fernando and C. P. Tsokos, "Dynamically Weighted Balanced Loss: Class Imbalanced Learning and Confidence Calibration of Deep Neural Networks", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 33, No. 7, pp. 2940-2951, Jul. 2022. <https://doi.org/10.1109/TNNLS.2020.3047335>.
- [59] M. Priyadharshini, V. Muruges, and O. Rybin, "Enhancing Alzheimer's disease classification with a transformer-based model using self-supervised learning", *Scientific Reports*, Vol. 16, Art. no. 3798, Jan. 2026. <https://doi.org/10.1038/s41598-025-33957-w>.
- [60] S. Bit, et al., "MRI-based mild cognitive impairment and Alzheimer's disease classification using an algorithm of combination of variational autoencoder and other machine learning classifiers", *Journal of Alzheimer's Disease Reports*, Vol. 8, No. 1, pp. 1434-1452, Oct. 2024. <https://doi.org/10.1177/25424823241290694>.
- [61] S. Wang, H. Zhang, and W. Kong, "Multimodal Classification of Alzheimer's Disease Using Longitudinal Data Analysis and Hypergraph Regularized Multi-Task Feature Selection", *Bioengineering*, Vol. 12, No 4, Art. no. 388, Apr. 2025. <https://doi.org/10.3390/bioengineering12040388>.
- [62] C. Molnar, G. König, B. Bischl, and G.

Casalicchio, "Model-agnostic Feature Importance and Effects with Dependent Features -- A Conditional Subgroup Approach", arXiv preprint arXiv:2006.04628, pp. 1-40, Jun. 2020. <https://doi.org/10.48550/arXiv.2006.04628>.

[63] 5.2. Permutation feature importance, https://scikit-learn.org/stable/modules/permutation_importance.html. [accessed: Jan. 04, 2026]

저자소개

송 정 아 (Jung-a Song)



2023년 3월 ~ 현재 : 부경대학교
데이터정보과학부
빅데이터융합전공 학사과정
관심분야 : 기계학습, 생물정보학,
데이터마이닝, 딥러닝

윤 태 현 (Taehyeon Yun)



2025년 2월 : 국립부경대학교
해양바이오신소재학과(이학사)
2025년 2월 : 국립부경대학교
빅데이터융합전공(공학사)
2025년 3월 ~ 현재 :
국립부경대학교 데이터공학과
석사과정

관심분야 : 기계학습, 생물정보학, 데이터마이닝, 딥러닝,
의료인공지능

하 지 환 (Jihwan Ha)



2020년 8월 : 연세대학교
컴퓨터과학과(공학박사)
2020년 9월 ~ 2021년 8월 :
하와이 암센터 포스닥 연구원
2021년 9월 ~ 2025년 8월 :
부경대학교 데이터정보과학부
빅데이터융합전공 조교수

2025년 9월 ~ 현재 : 부경대학교 데이터정보과학부
빅데이터융합전공 부교수

관심분야 : 기계학습, 생물정보학, 데이터마이닝, 딥러닝,
추천시스템