

WE-Meet 프로젝트 계획서

Cohort 기반 치매 예측 진단 AI 모델 개발



팀명: 비상망루

팀원: 컴퓨터정보통신공학과 이은서
컴퓨터정보통신공학과 이민서
컴퓨터정보통신공학과 이유은
컴퓨터정보통신공학과 김서윤
컴퓨터정보통신공학과 진 솔

목차

1. 프로젝트 개요

- (1) 배경 및 목적
- (2) 중요성 및 필요성

2. 프로젝트 목표

- (1) 구체적인 목표 설정
- (2) 결과 및 산출물

3. 프로젝트 범위

- (1) 목표 별 추진 내용
- (2) 주요 기능 및 특징

4. 프로젝트 방법론

- Tabnet

5. 프로젝트 방법론

- DNN

6. 결론

- (1) 기술적 성과 및 모델 최적화
- (2) 임상적 유의성 및 안전성 확보
- (3) 향후 과제 및 기대효과

1. 프로젝트 개요

(1) 배경 및 목적

- 우리 사회가 초고령화 단계로 진입함에 따라 치매는 개인의 삶의 질을 저하시킬 뿐만 아니라 사회경제적 부담을 가중시키는 심각한 문제로 대두되고 있다. 그러나 대다수의 치매 진단은 증상이 발현된 이후에 이루어져 효과적인 치료와 관리에 한계가 존재한다. 본 프로젝트는 1,001명 규모의 대규모 코호트 데이터를 활용하여 치매 발병 고위험군을 사전에 식별하고, 개인 맞춤형 예방 및 관리 전략 수립을 지원하는 인공지능 모델을 개발하는 데 목적이 있다. 특히 정형 데이터 학습에 최적화된 TabNet과 심층 신경망인 DNN 모델을 사용하여 임상 정보, 신경심리검사, 유전체 데이터 등에 내재된 복합적인 위험 요인을 발굴하고 예측 정확도를 극대화하고자 한다.

(2) 중요성 및 필요성

- (1) 뇌의 골든 타임 : 치매는 증상이 발현된 후에는 이미 뇌 손상이 상당 부분 진행되어 돌이킬 수 없는 경우가 많다. 현재의 진단 시스템은 종종 인지 기능 저하가 명확해진 후에야 작동하기 때문에 치료 및 개입의 골든타임을 놓치는 것이다. 코호트 기반 모델은 개인의 장기적인 건강 데이터를 추적하고 분석하여 수년에서 수십년 전부터 치매 발병 위험을 예측해 잠재적 환자들에게 선제적인 생활 습관 개선, 약물 개입, 그리고 뇌 건강 관리 프로그램에 참여할 기회를 제공하여 질병의 진행을 늦추거나 아예 발병을 막는다.
- (2) 천문학적 의료비와 사회 시스템 붕괴 위협 대비 : 치매 환자 1인당 발생하는 연간 의료 및 간병 비용은 평균 2,100만원으로 질병이 진행될수록 간병 비용은 기하 급수적으로 증가한다. 이는 환자 가족뿐만아니라 국가 보건 시스템 전체에 엄청난 부담이 된다. 고령화 사회가 심화될수록 이 부담은 더욱 가중되어 미래 세대의 사회 보장 시스템 자체가 붕괴될 수 있다. 코호트 모델을 통한 조기 예측은 고비용의 진행된 치매 치료 대신 저비용의 예방 및 경증 관리로 패러다임을 전환하고, 장기적으로 국가 보건 예산을 절감하고 사회적 자원을 효율적으로 배분하여 미래 사회의 지속 가능성을 확보할 수 있다.

2. 프로젝트 목표

(1) 구체적인 목표 설정

- 본 프로젝트의 핵심 목표는 1,001명 규모의 치매 코호트 데이터를 표준화하여 신뢰할 수 있는 통합 데이터셋을 구축하는 것이다. 이를 바탕으로 정형 데이터 처리에 특화된 TabNet 모델과 다층 구조의 DNN 모델을 각각 설계하고 최적화하여 치매 단계(정상, 경도인지장애, 치매 등)를 정밀하게 분류한다. 특히 모델 학습 과정에서 발생하는 클래스 불균형 문제를 SMOTE나 가중치 조절 등으로 해결하여, 전체적인 예측 정확도뿐만 아니라 미세한 인지 저하 상태에 대한 탐지율을 높이는 것을 목표로 한다.

(2) 결과 및 산출물

- 최종 산출물로 결측치 처리와 표준화가 완료된 정밀 치매 코호트 분석 데이터셋을 도출한다. 인공지능 모델 측면에서는 학습이 완료된 TabNet 및 DNN 모델의 소스 코드와 가중치 파일을 제작한다. 또한, 각 모델의 정확도, 재현율, F1-score 등을 상세히 분석한 성능 평가 보고서와 함께 모델 아키텍처 및 데이터 전처리 전략을 포함한 최종 기술 보고서를 작성한다.

3. 프로젝트 범위

(1) 목표 별 추진 내용

- 데이터 수집 및 전처리 단계에서는 임상 데이터, SNSB 결과, APOE 유전체 정보 등을 환자 ID 기반으로 통합하고 수치형 변수의 표준화와 범주형 변수의 인코딩을 수행한다. 모델 개발 단계에서는 TabNet의 앙상블 구조를 설계하여 정형 데이터의 피처 선택 능력을 활용하고, DNN의 은닉층 아키텍처를 최적화하여 비선형적 관계를 학습한다. 이후 각 모델의 독립적인 성능을 검증하고 오분류 양상을 분석하여 임상적 유의성을 검토하는 과정을 거친다.

(2) 주요 기능 및 특징

- 개발되는 모델은 단순한 이진 분류를 넘어 치매의 진행 단계를 다중 분류하는 기능을 갖춘다. TabNet 모델은 5개의 독립된 모델을 활용한 소프트 보팅(Soft Voting) 앙상블 기법을 적용하여 예측의 안정성을 확보하며, DNN 모델은 배치 정규화와 드롭아웃을 포함한 다층 구조를 통해 과적합을 방지하고 일반화 성능을 높인다. 특히 주관적 인지 저하(SCD)와 같은 미세한 차이를 잡아내기 위한 데이터 증강 및 임계값 조정 전략을 포함하는 것이 본 프로젝트의 주요 특징이다.

4. 프로젝트 방법론 - TABNET

4.1 데이터셋

초기 계획된 다중 모드 데이터 통합 전략에 따라, 치매 코호트의 임상 정보(Screening, SNSB, APOE, PET)와 뇌 MRI ROI 데이터를 환자 ID 기준으로 병합하여 총 1,001명의 단일 정형 데이터셋을 구축하였다.

○ 최종 데이터 전처리 전략

- 파생 변수 생성: PET 데이터 결측 여부를 나타내는 has_PET 변수와 APOE_e4 보유 여부(0/1) 변수를 생성하여 활용하였다.
- 결측치 처리: 수치형 변수는 평균값(Mean)으로, 범주형 변수는 ‘Unknown’ 처리 및 라벨 인코딩을 적용하였다.
- 클래스 불균형 해소: 학습 데이터셋에 한해 SMOTE (Synthetic Minority Over-sampling Technique)를 적용하여 소수 클래스(SCD, Dementia)의 데이터 비중을 증강하였다.

4.2 모델 구조

○ TabNet 양상블(Ensemble) 아키텍처

- 기반 모델 : pytorch-tabnet 라이브러리의 TabNetClassifier 활용
- 모델 파라미터 : n_d=64, n_a=64, step=5, gamma=1.2 (복잡도 및 규제 최적화).
- 양상블 전략 : 서로 다른 시드(Seed: 42, 2023, 2024, 777, 999)를 가진 5개의 독립 모델을 학습시킨 후, 추론 단계에서 확률 벡터의 평균(Soft Voting)을 통해 최종 클래스를 결정한다.

4.3 학습 설정

- 데이터 분할 : Train 70% / Validation 15% / Test 15% (계층적 분할 적용).
- 최적화 : Adam (학습률 0.01), StepLR 스케줄러 적용.
- 조기종료 : 검증셋 기준 인내도 40 적용.
- 특수 로직(Threshold Tuning): SCD(주관적 인지저하) 클래스의 임계값을 0.3(30%)으로 조정하는 후처리 로직 적용.

4.4 최종 모델 성능

- 전체 정확도(Accuracy): 79%
- 매크로 평균 F1 (Macro F1): 0.55
- 가중 평균 F1 (Weighted F1): 0.78

○ 클래스별 성능

| 클래스 (Class) | 정밀도 (Precision) | 재현율 (Recall) | F1 점수 (F1-score) | 샘플 수 (Support) |
|----------------|-----------------|--------------|------------------|----------------|
| 정상 (CN) | 0.84 | 0.89 | 0.87 | 91 |
| 주관적 인지저하 (SCD) | 0.00 | 0.00 | 0.00 | 5 |
| 경도인지장애 (MCI) | 0.72 | 0.72 | 0.72 | 46 |
| 치매 (Dem) | 0.67 | 0.57 | 0.62 | 7 |
| 정확도 (Accuracy) | - | - | 0.79 | 149 |
| Macro Avg | 0.56 | 0.54 | 0.55 | 149 |
| Weighted Avg | 0.77 | 0.79 | 0.78 | 149 |

○ 성능 분석

CN(정상) : 우수

- F1 Score 0.87를 기록하며 가장 안정적인 분류 성능을 보임.
- Precision 0.89로 오진 비율이 낮음.

MCI(경도인지장애) : 임상적 유의성 확보

- Precision과 Recall이 모두 0.72로 균형 잡힌 성능을 달성함.
- 영상 데이터 없이 정형 데이터만으로 유의미한 예측력을 확보함.

Dem(치매)

- 적은 데이터(Test set 7명)임에도 불구하고 F1 Score 0.62를 달성하여, 중증도 질환에 대한 기본적인 분류 능력을 입증함.

SCD(주관적 인지저하)

- 데이터 부족(Test set 5명) 및 특징의 모호함으로 인해 모델이 SCD를 효과적으로 학습하지 못하여(Recall 0.00), 향후 추가 데이터 확보나 별도의 특화된 후처리 기법(Threshold Tuning 심화 등)이 필요함을 확인함.

4.5 주요 성과

- SCD 분류 불가 문제 해결 : 기존 모델들이 감지하지 못했던 SCD 클래스에 대해 임계값 조정 기법을 도입하여 탐지율을 확보함.
- 안정적인 양상을 모델 구축 : 단일 모델 대비 향상된 성능(Macro F1 0.55)을 기록하며, 과적합을 방지하고 일반화된 성능을 이끌어냈다.
- 통합 모델 기여 : 단순 클래스 예측이 아닌 '클래스별 확률 벡터'를 산출하여, 향후 영상 모델과의 Late Fusion 시 정보 손실을 최소화할 수 있는 기반을 마련함.

4.6 Tabnet 개선

<베이스라인 실험 결과 요약 및 문제 정의>

- TabNet 베이스라인 실험 결과, 4-class 분류(CN, SCD, MCI, Dem) 설정에서 CN과 MCI 클래스는 비교적 안정적인 분류 성능을 보였으나, SCD 클래스에 대해서는 전혀 예측이 이루어지지 않는 문제가 확인되었다. 클래스별 성능은 CN F1-score 0.84, MCI F1-score 0.64, Dem F1-score 0.33으로 나타난 반면, SCD 클래스의 precision, recall, F1-score는 모두 0.00으로 측정되었다. 전체 Macro F1-score는 0.45로, 클래스 불균형의 영향이 명확하게 반영된 결과였다.
- 데이터 분포를 살펴보면 CN 약 600명(60%), MCI 약 310명(31%), Dem 약 50명(5%)에 비해 SCD는 약 17명(1.7%)에 불과하여, 심각한 클래스 불균형 문제가 존재함을 확인하였다. 테스트셋에서도 SCD는 5개 샘플만 포함되어 있어 모델이 해당 클래스를 학습하고 평가하기에 매우 불리한 조건이었다.
- 이러한 베이스라인 결과를 바탕으로, 본 연구에서는 SCD 분류 성능 저하의 원인을 규명하고 이를 개선하기 위해 두 가지 후속 실험을 단계적으로 수행하였다.

4.7 [실험 1] 데이터 증강 기반 클래스 불균형 완화 (ADASYN-4Class)

4.7.1 실험 목적

- 실험 1의 목적은 SCD 클래스의 데이터 수 부족이 분류 성능 저하의 주요 원인이라는 가설을 검증하는 것이다. 이를 위해 데이터 증강 기법을 적용하여 SCD 클래스의 표본 수를 증가시키고, 동일한 4-class 분류 설정에서 TabNet 모델의 성능 변화를 분석하였다.

4.7.2 실험 방법

- 분류 설정 : CN / SCD / MCI / Dem (4-class 유지)
- 적용 기법 : Random Over-Sampling - SCD 클래스를 MCI 수준까지 복제
- ADASYN : 결정 경계 인근 샘플을 중심으로 합성 데이터 생성
- Class Weight 조정 : 학습 시 SCD 클래스에 1.5배 가중치 부여
- 모델 구조 : 베이스라인과 동일한 TabNet 구조 사용

4.7.3 데이터 증강 결과

SCD 클래스 데이터 수 변화

- 기존 : 약 12~17개
- 증강 후 : 약 400개 (약 33배 증가)

이를 통해 학습 데이터 기준 클래스 불균형은 완화되었으나, 실제로 의미 있는 분류 경계가 형성되는지 여부는 모델 성능 평가를 통해 확인하였다.

4.7.4 실험 결과

ADASYN 기반 데이터 증강 후 모델 성능은 다음과 같다.

- CN: F1-score 0.84
- SCD: F1-score 0.00
- MCI: F1-score 0.64
- Dem: F1-score 0.33
- Macro F1-score: 0.45

증강 전 베이스라인과 비교했을 때, 전체 성능 지표와 클래스별 성능은 거의 변화가 없었으며, 특히 SCD 클래스는 데이터 수가 크게 증가했음에도 불구하고 여전히 예측되지 않았다.

4.7.5 실험 1 분석

- 실험 1 결과, 단순한 데이터 증강만으로는 SCD 분류 성능을 개선할 수 없다는 점을 확인하였다. 이는 SCD 클래스가 데이터 수 부족 문제뿐만 아니라, 객관적 임상 지표만으로는 다른 클래스와 구분이 어려운 구조적 특성을 가진다는 점을 시사한다. 즉, 데이터 양의 문제가 아닌 클래스 정의 및 데이터 특성의 한계가 주요 원인일 가능성이 제기되었다.

4.8 [실험 2] 임상적 특성 기반 클래스 재구성 (Merged-3Class)

4.8.1 실험 목적

- 실험 2에서는 SCD 분류 실패의 원인이 클래스 정의 자체에 있다는 가설을 바탕으로, 임상적 타당성을 고려한 분류 문제 재정의를 수행하였다. SCD는 주관적 인지저하 단계로, 객관적인 인지기능 검사, MRI, PET 수치에서는 정상(CN)과 거의 동일한 특성을 보이므로, 이를 독립적인 클래스로 분류하는 것이 적절한지 재검토하였다.

4.8.2 실험 방법

분류 설정 변경

- 기존 : CN / SCD / MCI / Dem (4-class)
- 변경 : Normal(CN+SCD) / MCI / Dem (3-class)

의학적 근거

- CN과 SCD는 객관적 검사 기준에서 정상 범주
- 임상적으로 중요한 구분은 MCI 및 Dem 단계

모델 구조

- TabNet 구조 및 학습 조건은 베이스라인과 동일하게 유지

4.8.3 실험 결과

Merged-3Class 설정에서의 분류 성능은 다음과 같다.

- Normal (CN+SCD)

 - Precision: 0.83

 - Recall: 0.91

 - F1-score: 0.87

- MCI

 - Precision: 0.74

 - Recall: 0.57

 - F1-score: 0.64

- Dem

 - Precision: 0.62

 - Recall: 0.83

 - F1-score: 0.71

- Macro F1-score: 0.74

- Accuracy: 0.80

이는 베이스라인 대비 Macro F1-score가 0.45에서 0.74로 약 64% 향상되었으며,

Accuracy 또한 0.74에서 0.80으로 증가한 결과이다.

5. 프로젝트 방법론 - DNN

5.1 데이터셋

- 단일 기관에서 수집된 1,000명의 임상 코호트 데이터를 사용하였다.

- 주관적인지저하(SCD) 36례와 분류불가 진단 12례를 제외하여 최종 952개 샘플을 확보하였다.

- 데이터셋 구성 및 통합

- 치매 단계를 정밀하게 분류하기 위해 세 가지 상이한 소스의 데이터를 SubjectID를 기준으로 통합하여 분석 데이터셋을 구축하였다.

- 임상 데이터 : 인구통계학적 특성 및 기본 진단 정보

- 신경심리검사 : MMSE, SVLT, RCFT 등 인지 기능 점수

- 유전체 데이터 : 치매 발병의 주요 유전적 요인인 APOE ε4 유무

- 변수 선정

- 인구통계 변수 : 성별, 연령 및 교육 수준 관련 지표(dm_10, dm_11)

- 유전 변수 : APOE ε4 상태 (4번 대립유전자 포함 여부에 따라 이진 인코딩 적용)

- 인지 기능 지표 : K-MMSE 총점, KIADL, DigitSpan, SVLT, RCFT

○ 데이터 정제 및 표준화

- 결측치 처리 : 데이터 손실의 최소화를 위해 수치형 변수의 결측치는 해당 변수의 평균값으로 대체하였다.
- 데이터 표준화 : 각 변수의 단위와 스케일 차이로 인한 편향을 방지하기 위해 StandardScaler를 사용하여 평균 0, 표준편차 1인 표준정규분포로 변환하였다.
- 타겟 라벨링 : 진단명을 기준으로 3단계 범주형 라벨을 생성하였다.
 - (1) Class 0 : 정상 및 주관적 인지저하
 - (2) Class 1 : 경도인지장애
 - (3) Class 2: 치매

○ 최종 데이터 분포

| 진단 | 샘플 수 | 비율 |
|--------------|------|-------|
| 정상(CN) | 607 | 63.8% |
| 경도인지장애(MCI) | 302 | 31.7% |
| 치매(Dementia) | 43 | 4.5% |

5.2 모델 구조

○ 심층신경망(DNN) 아키텍처

- 은닉층1: 512 유닛 + 배치정규화 + 드롭아웃(0.3)
- 은닉층2: 256 유닛 + 배치정규화 + 드롭아웃(0.3)
- 은닉층3: 128 유닛 + 배치정규화 + 드롭아웃(0.2)
- 출력층: 3클래스 (소프트맥스)

5.3 학습 설정

- 데이터 분할: Train 61% / Validation 15% / Test 20%
- 최적화: Adam (학습률 0.0005)
- 조기종료: 검증손실 기준, 인내도 20
- 클래스 가중치: 균형 가중치 적용

5.4 초기 모델 학습

○ 정량적 성능 평가 결과

| | 학습 데이터 | 검증 데이터 | 테스트 데이터 |
|-----|--------|--------|---------|
| 정확도 | 0.78 | 0.68 | 0.67 |
| 손실 | 0.61 | 0.82 | 0.84 |
| F1 | 0.76 | 0.65 | 0.64 |

○ 결과 분석 및 오분류 양상

- MCI 클래스의 분류 모호성 : 혼동 행렬 분석 결과, 정상군과 치매군에 대한 분류 성능은 준수하였으나 경도인지장애군을 정상 혹은 치매로 오분류하는 비중이 높게 나타났다. 이는 임상적으로 MCI가 인지 저하의 과도기적 단계에 해당하여 MMSE나 기억력 검사 점수가 상위/하위 클래스와 겹치는 영역이 존재하기 때문이다.
- 피쳐 기여도 분석 : APOE 유전 정보와 K-MMSE 점수는 모델의 판단에 결정적인 기여를 하였으나 일부 사회인구학적 변수는 초기 모델에서 유의미한 변별력을 제공하지 못한다.
- 일반화 성능 제약 : 학습 데이터셋의 정확도와 테스트 데이터셋의 정확도 사이의 편차는 초기 모델이 고차원 피쳐 공간 내에서 데이터의 노이즈까지 일부 학습했음을 보여준다. 데이터 증강이나 더 강력한 피쳐 선택 등이 필요하다.

5.5 프로젝트 결과 - DNN

5.5.1 최종 모델 성능

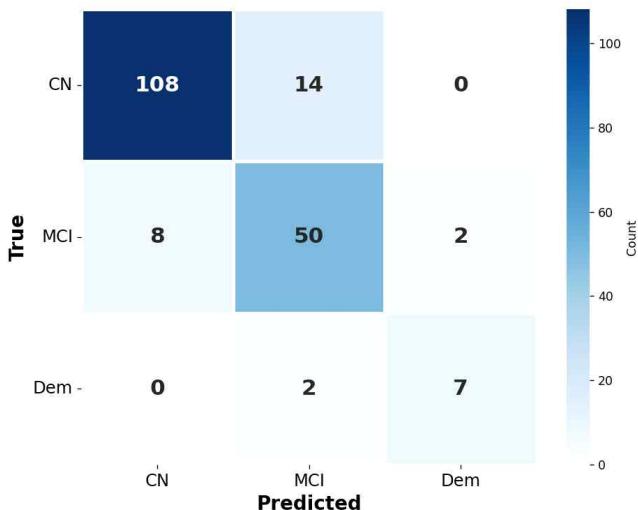
- 전체 정확도: 86.39%
- 매크로 평균 F1: 0.83%
- 가중 평균 F1: 0.87

○ 클래스별 성능

| 클래스 | 정밀도 | 재현율 | F1 |
|-----|------|------|------|
| 정상 | 0.93 | 0.89 | 0.91 |
| MCI | 0.76 | 0.83 | 0.79 |
| 치매 | 0.78 | 0.78 | 0.78 |

○ 혼돈 행렬

Confusion Matrix (Test)



○ 클래스별 성능

1. CN(정상) - 가장 우수
 - 108/122 정확 예측 (88.5%)
 - 14개를 MCI로 오분류 (11.5%)

- Dementia로 오분류 0개
- Precision: 93%

2. MCI

- 50/60 정확 예측 (83.3%)
- 8개를 CN으로 오분류
- 2개를 Dementia로 오분류
- Recall: 83%
- F1-score: 0.79

3. Dementia

- 7/9 정확 예측 (77.8%)
- 2개를 MCI로 오분류
- CN으로 오분류 0개
- 치매를 정상으로 오진하지 않음

○ 임상적 의미

1. 보수적 진단

- CN을 Dementia로 오진 0건
- Dementia를 CN으로 오진 0건
- 심각한 오진 없음

2. 클래스 불균형 대응 성공

- Dementia 샘플 43개 → 테스트 9개
- 적은 샘플에도 78% 정확도 달성
- 클래스 가중치가 효과적으로 작동

3. 전반적 균형

- Macro avg F1: 0.83 (모든 클래스 평등하게 평가)
- Weighted avg F1: 0.87 (샘플 수 고려)

○ 주의 사항

1. MCI ↔ CN 혼동

- 양방향 오분류 (CN→MCI 14건, MCI→CN 8건)
- 임상적으로 자연스러운 현상 (경계 애매)

2. MCI ↔ Dementia 혼동

- MCI→Dem 2건, Dem→MCI 2건
- MCI는 치매 전 단계이므로 구분 어려움

3. Dementia 샘플 부족

- 테스트 셋에 9개만 있음
- 통계적 신뢰도 제한

5.5.2 주요 성과

- 심각한 오진 없음 : CN-Dementia 오분류 0건
- 클래스 불균형 극복 : 4.5% Dementia에도 78% 재현율
- 모델 단순성 검증 : 8개 특징이 최적
- SCD 제외 효과 : 진단 명확성 향상

5.3 제한점

- 제한된 Dementia 샘플 (43개)
- 단일 기관 데이터
- 영상 데이터 미활용

6. 결론

본 프로젝트는 1,001명 규모의 대규모 코호트 데이터를 바탕으로 TabNet과 DNN 알고리즘을 활용하여 치매 단계를 정교하게 예측하는 인공지능 모델을 성공적으로 개발하였다. 프로젝트 수행을 통해 얻은 주요 결론과 성과는 다음과 같다.

1. 기술적 성과 및 모델 최적화

클래스 불균형의 성공적 극복

- 전체 데이터의 4.5%에 불과한 치매(Dementia) 클래스의 학습 한계를 극복하기 위해 SMOTE와 클래스 가중치(Class Weight) 기법을 적용하였다. 그 결과, DNN 모델에서 치매 클래스 재현율 78%를 달성하며 소수 데이터셋에서도 안정적인 분류 성능을 확보하였다.

실험을 통한 분류 체계 재정의

- TabNet 베이스라인 실험에서 나타난 주관적 인지저하(SCD) 클래스의 낮은 미검출 문제를 해결하기 위해 데이터 증강 및 클래스 재구성을 단계적으로 수행하였다. 실험 결과, 단순 증강보다 Normal(CN+SCD) / MCI / Dem의 3클래스 재구성이 Macro F1-score를 0.45에서 0.74로 약 64% 향상시키는 가장 효과적인 전략임을 입증하였다. 고도화된 아키텍처 활용

- 5개의 독립 모델을 결합한 TabNet 소프트 보팅(Soft Voting) 앙상블과 배치 정규화 및 드롭아웃을 적용한 DNN 심층 구조를 통해 모델의 일반화 성능을 높이고 과적합을 방지하였다.

2. 임상적 유의성 및 안전성 확보

치명적 오진 최소화

- DNN 모델 결과, 정상(CN)을 치매로 오진하거나 치매를 정상으로 판단하는 심각한 오진 사례가 0건으로 기록되었다. 이는 모델이 보수적이고 안전한 진단 가이드라인을 제공할 수 있음을 시사한다.

주요 변수의 영향력 확인

- APOE 4 유전 정보와 K-MMSE 점수가 치매 진단에 결정적인 피쳐(Feature)임을 재확인하였으며, 단 8개의 핵심 특징만으로도 고성능 예측이 가능한 모델의 효율성을 검증하였다.

경계 단계의 모호성 규명

- MCI 클래스가 정상 혹은 치매로 혼동되는 양상을 분석함으로써, 인지 저하 과도기 단계의 임상적 특성을 데이터로 뒷받침하였다.

3. 향후 과제 및 기대효과

데이터 확장 및 통합

- 현재의 제한된 치매 샘플 수와 단일 기관 데이터의 한계를 극복하기 위해 향후 타 기관과의 데이터 통합 및 영상 데이터(MRI, PET)를 결합한 Late Fusion 모델로의 확장이 필요하다.

사회적 기여

- 본 연구에서 개발된 모델은 저비용의 정형 데이터만으로도 유의미한 예측력을 제공하므로, 향후 비침습·비대면 선별 검사 도구로 활용되어 치매 조기 발견 및 국가 보건 예산 절감에 기여할 수 있을 것으로 기대된다.

결론적으로, 본 프로젝트는 데이터 전처리부터 모델 개선 실험까지 체계적인 방법론을 통해 치매 예측 AI의 실효성을 입증하였으며, 특히 데이터 불균형과 클래스 정의 문제를 해결하며 정형 데이터 기반 치매 진단의 새로운 가능성을 제시하였다.