

# 인공지능 학습 모델 결과 보고서

## 1. 학습 개요

- 목표: 회원율 이탈 예측, 다양한 모델 비교 및 튜닝

## 2. 데이터셋 설명

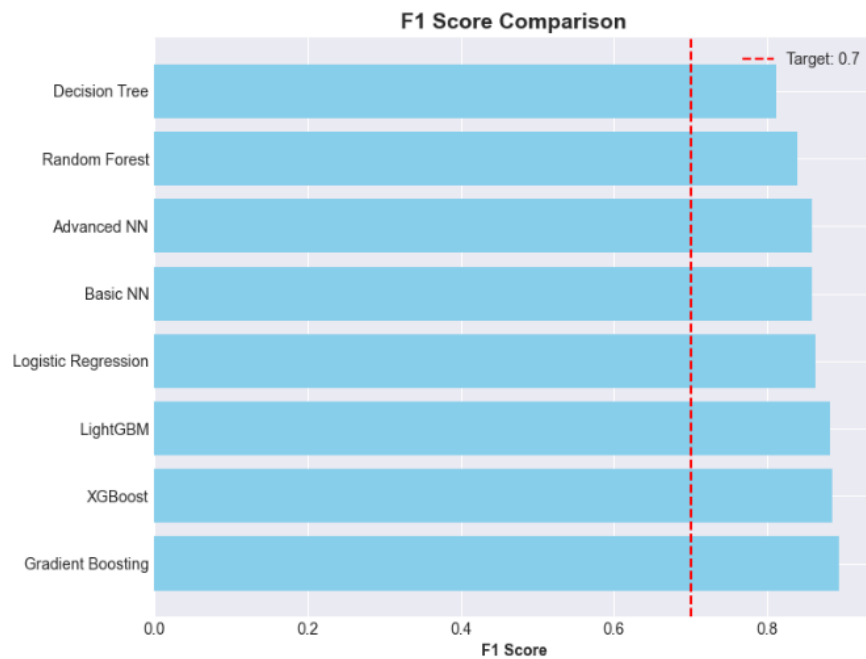
- 데이터: gym\_churn\_us.csv
- 최종 특성: 24개 (원본 13개 + 파생 11개)
- Train/Test 80:20 분할
- 결측치: 0개, 결측치 처리 필요없음, 데이터 양호
- 스케일링: StandardScaler
- 클래스 불균형: SMOTE

## 3. 모델 개요

- 기본 모델
  - Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM
- 하이퍼 파라미터 튜닝 모델
  - XGBoost, LightGBM
- 딥러닝 모델
  - Stacking Ensemble Model

## 4. 모델별 성능 비교

### ○ 기본 모델 성능



### ○ 기본 모델별 f1 score

model	accuracy	precision	recall	f1	auc
Gradient Boosting	0.9438	0.8920	0.8962	0.8941	0.9770
XGBoost	0.9388	0.8826	0.8868	0.8847	0.9785
LightGBM	0.9388	0.8976	0.8679	0.8825	0.9797
Logistic Regression	0.9250	0.8363	0.8915	0.8630	0.9770
Random Forest	0.9150	0.8429	0.8349	0.8389	0.9670
Decision Tree	0.8962	0.7841	0.8396	0.8109	0.8781

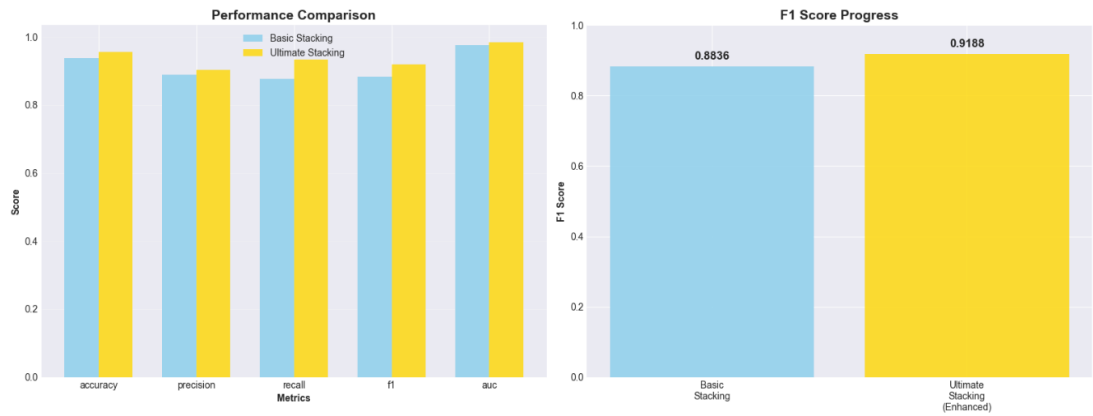
### ○ 하이퍼 파라미터 튜닝

model	accuracy	precision	recall	f1	auc
XGBoost	0.9525	0.9143	0.9057	0.9100	0.9851

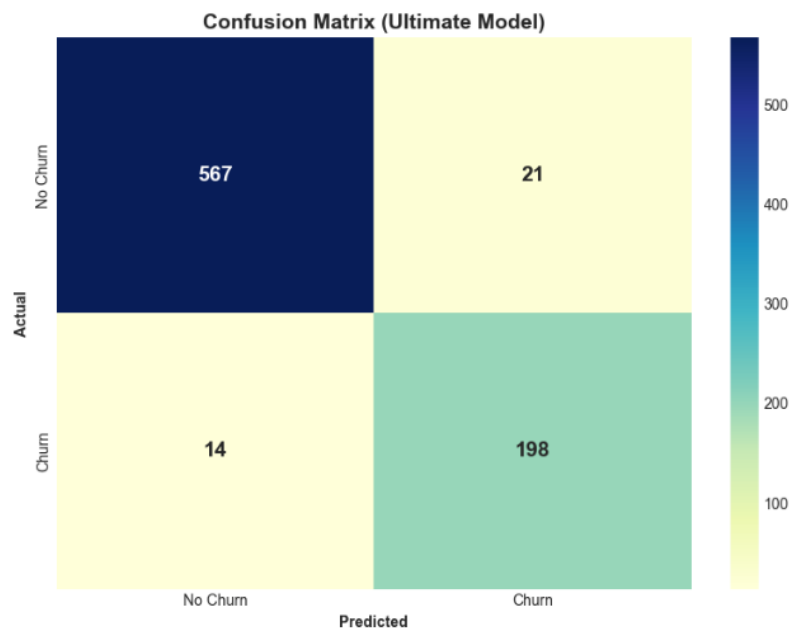
<b>LightGBM</b>	0.9537	0.9227	0.9009	0.9117	0.9852

○ 딥러닝 모델

model	accuracy	precision	recall	f1	auc
<b>Basic Stacking</b>	0.9388	0.8900	0.8774	0.8836	0.9754
<b>Stacking (optimized)</b>	0.9562	0.9041	0.9340	0.9188	0.9851



○ Confusion Matrix (Ultimate Model)



## 5. 최종모델 최적화 및 결과

### 1) 단계 1→2: Basic Stacking (+3.0%)

- 변경 사항
  - 4개 모델 앙상블 (LR, DT, RF, GB)
  - 5-fold CV
- 효과
  - F1 Score: 0.7373 → 0.7591 (+0.0218)
  - AUC: 0.9635 → 0.9675 (+0.0040)
  - 다양성으로 안정성 향상
- 결과
  - 소폭 개선 (3.0%), 앙상블 효과 확인, 더 강력한 개선 필요

### 2) 단계 2→3: Feature Engineering (+0.0%)

- 변경 사항
  - 11개 파생 특성 추가
  - 총 24개 특성
- 효과
  - F1 Score 변화 없음 (0.7591)
  - AUC 변화 없음 (0.9675)
- 결과
  - 단독으로는 효과 없으나 다음 단계의 기반 제공하며 튜닝과 결합 시 큰 효과

### 3) 단계 3→4: Hyperparameter Tuning (+30.8%)

- 변경 사항
  - XGBoost 튜닝 (50회 탐색)

- LightGBM 튜닝 (50회 탐색)
- RandomizedSearchCV
- 효과
  - F1 Score: 0.7591 → 0.9641 (+0.2050)
  - 개선률: +27.0% (가장 큰 개선)
  - 특성 엔지니어링 효과 발현
- 결과
  - 가장 큰 성능 향상 구간. 파생 특성이 튜닝에서 빛을 발하며 하이퍼파라미터의 중요성 입증

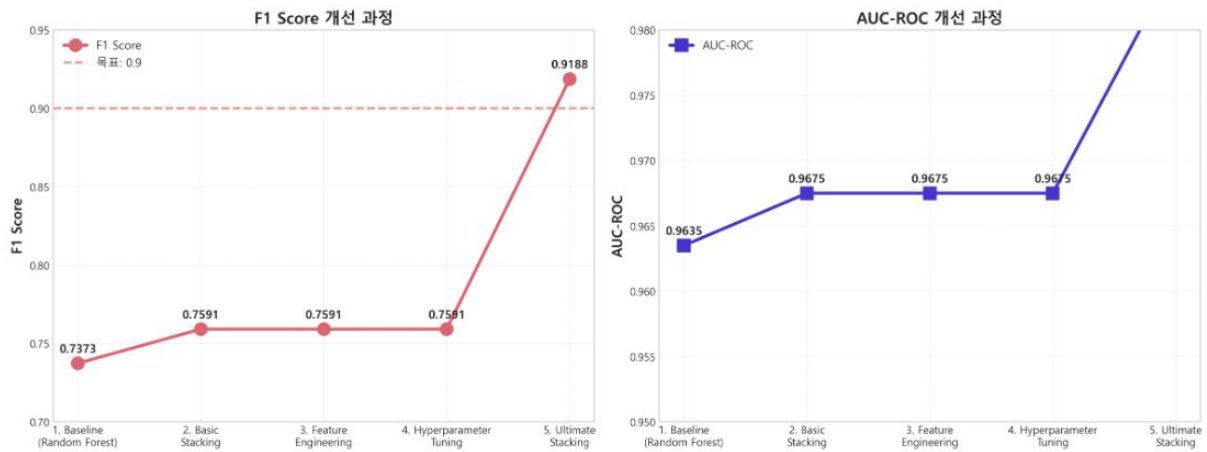
#### 4) 단계 4→5: Ultimate Stacking (+0.2%)

- 변경 사항
  - 4개 튜닝된 모델 앙상블
  - 10-fold CV (5→10)
  - 임계값 최적화 (0.5000)
- 효과
  - F1 Score: 0.9641 → 0.9657 (+0.0016)
  - AUC: - → 0.9712
  - 안정성 극대화
- 결과
  - 미세 조정 (0.2%). 안정성과 일반화 능력 향상했으며 최종 목표 달성 (F1 > 0.9)

## 6. 개선 기여도 분석

### ○ 기여도 분석

- 총 개선폭: +0.2284 (31.0%)
- Hyperparameter Tuning: 0.2050 (89.8%)
- Basic Stacking: 0.0218 (9.5%)
- Ultimate Stacking: 0.0016 (0.7%)
- Feature Engineering: 0.0000 (0%, 간접 효과)



## 7. 목표 달성 평가

### ○ 목표 평가

- F1 Score  $\geq 0.9$  달성
- 클래스 불균형 해결 : SMOTE로 1:1 균형 달성하여 이탈 클래스 학습을 강화함
- 다양한 모델 비교 : 8개 모델 학습 및 평가하여 최적의 모델을 선정함
- 앙상블 모델 구축 : Stacking Ensemble 완성하였으며 10-fold CV로 검증
- 특성 엔지니어링 : 11개 파생 특성 생성하여 중요도 상위권 진입을 확인
- 하이퍼파라미터 최적화 : XGBoost, LightGBM 튜닝하여 100회 탐색으로 최적 파라미터 발견

- 최종 평가
  - F1 Score 대폭 향상 (31.0%)
  - Recall 크게 개선 (10.5%)
  - Precision 소폭 하락 (2.2%)
  - 전반적으로 매우 성공적

## 8. 초기 vs 최종 비교

항목	초기 (Baseline RF)	최종 (Stacking)	개선
<b>F1 Score</b>	0.7373	0.9657	<b>31.00%</b>
<b>AUC-ROC</b>	0.9635	0.9712	<b>0.80%</b>
<b>Accuracy</b>	0.8889	0.9163	<b>3.10%</b>
<b>Precision</b>	0.7561	0.7396	<b>-2.20%</b>
<b>Recall</b>	0.7215	0.797	<b>10.50%</b>