

# 강의계획서 25주차

## [ 강의계획서 ]

항목	내용
과목명	공공데이터를 활용한 AI 이해와 예측 모델 구축
교·강사명	
연도/학기	2026

## 25차시 강의내용

차시	내용	구분	과제
25	<p>1) 차시명: 모델 저장과 실무 배포 준비</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"><li>- pickle/joblib을 사용하여 학습된 모델을 저장하고 불러온다</li><li>- 실무 배포 시 고려해야 할 핵심 사항을 이해한다</li><li>- 지속적인 모델 운영의 필요성과 다음 학습 방향을 파악한다</li></ul> <p>3) 학습내용:</p> <ul style="list-style-type: none"><li>- 모델 저장의 필요성 (매번 재학습 vs 저장된 모델 사용)</li><li>- pickle vs joblib 비교 (sklearn 모델에는 joblib 권장)</li><li>- 저장된 모델을 FastAPI/Streamlit에서 불러오기</li><li>- 실무 배포 체크리스트 (입력 검증, 버전 관리, 로깅)</li></ul>	실습	최종과제

# 세부 학습 내용

## 1. 모델 저장의 필요성

- 매번 모델을 재학습하는 것은 시간과 자원 낭비
- 학습된 모델을 저장하면 즉시 예측 서비스 제공 가능
- 모델 버전 관리 및 롤백 가능

## 2. pickle vs joblib 비교

항목	pickle	joblib
표준 라이브러리	O	X (별도 설치)
NumPy 배열 처리	일반적	최적화됨
대용량 모델	느림	빠름
sklearn 권장	-	O

## 3. 실습 코드 예시

```
# 모델 저장 (버전 정보 포함)
import joblib
from datetime import datetime

model_version = datetime.now().strftime("%Y%m%d_%H%M")
model_filename = f"quality_model_v{model_version}.joblib"
joblib.dump(model, model_filename)

# 모델 메타정보 저장
model_info = {
    'version': model_version,
    'accuracy': model.score(X_train, y_train),
    'features': ['temperature', 'pressure', 'humidity', 'vibration'],
    'created_at': datetime.now().isoformat(),
    'model_type': 'RandomForestClassifier',
    'n_estimators': 100
}

# 모델 불러오기 및 예측
```

```
loaded_model = joblib.load(model_filename)
prediction = loaded_model.predict(new_data)
```

4. 실무 배포 체크리스트

항목	설명
입력 데이터 검증	타입, 범위, 결측치 확인
모델 버전 관리	파일명에 날짜/버전 포함
예측 로깅	입력, 출력, 타임스탬프 기록
에러 처리	예외 상황 대비 (try-except)
성능 모니터링	정확도 변화 추적

최종 과제 안내

주제: 제조 품질 예측 웹 서비스 구축 (EDA - 모델링 - 서비스화)

평가 기준:

항목	비중	내용
문제 정의	10%	데이터 기반 문제 설정의 명확성
데이터 분석	20%	전처리 및 EDA의 적절성
모델링	30%	모델 선택, 학습, 튜닝의 타당성
모델 평가	20%	평가 지표 해석 및 개선 방향 도출
서비스 구현	10%	Streamlit/FastAPI 활용
보고서	10%	분석 결과 정리 및 비즈니스 해석

본 강의계획서는 AI 기초체력훈련(Pre AI-Campus) 심사 기준에 맞추어 작성되었습니다.