

[강의계획서 초안]

oswar_000

2025/12/16 04:31

초안이므로 강의 이름은 추후.. 진행

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

강의개요	본 과정은 AI S/W 개발자 양성을 위한 기초체력과정으로, 비전공자와 입문자가 AI 기술 개발 분야에 진입하기 위해 필요한 핵심 역량을 체계적으로 함양함. Python 프로그래밍 환경 구축부터 시작하여 공공데이터 활용, 통계 기반 데이터 분석, 머신러닝·딥러닝 모델 구현, 그리고 AI 모델의 API 연동 및 서비스화까지 실무 중심의 학습을 진행함. '문제정의→분석→모델링→평가' 단계별 학습구조를 통해 AI S/W 개발 과정 전반을 이해하고, 향후 심화 과정(KDT AI Campus)에 원활히 적용할 수 있는 기초 역량을 확보함.
사전필요지식	특별한 사전 지식 불필요. 컴퓨터 기본 활용 능력(파일 관리, 인터넷 검색) 보유 시 수강 가능함. 프로그래밍 경험이 있으면 학습에 도움이 되나 필수 요건은 아님.
훈련핵심역량	Python 프로그래밍 및 데이터 처리 기초, 기술통계 및 가설검정 기반 데이터 분석, 머신러닝·딥러닝 모델 설계 및 구현, CNN/RNN 등 주요 신경망 아키텍처 활용, 공개 데이터셋(AI 허브, Kaggle 등) 활용 능력, AI 모델의 웹 서비스화(Streamlit, FastAPI) 기초
사후취득지식	AI 작동원리와 관련 수리내용(확률, 통계, 선형대수) 기초 이해, 공공데이터·기업데이터를 활용한 문제 해결 능력, AI S/W 개발 과정(문제정의→분석→모델링→평가) 전반 이해, KDT(AI Campus) 과정 적응을 위한 기초 역량 확보
주교재	[주교재] <ul style="list-style-type: none">- 강의별 제공 교안, 실습 코드(Jupyter Notebook)- 오렐리언 제롬, 핸즈온 머신러닝(3판), 2023, 한빛미디어- 사이트 고기, 밑바닥부터 시작하는 딥러닝 시리즈, 2017~2024, 한빛미디어 [부교재 및 참고자료] <ul style="list-style-type: none">- 박해선, 혼자 공부하는 머신러닝+딥러닝, 2020, 한빛미디어- AI 허브(aihub.or.kr), Kaggle(kaggle.com), UCI ML Repository 공개 데이터셋

※ 실습과제: 차시별로 해당하는 실습 과제 제출(해당 차시 학습 후 풀이 가능)

※ 종합 프로젝트: 전 차시 학습 내용을 종합한 최종 프로젝트 제출(AI 모델 구축 및 서비스화)

1	1) 차시명: AI 개발 환경 구축 및 과정 오리엔테이션 2) 학습목표: <ul style="list-style-type: none">- AI 기초체력과정의 전체 구조와 학습 목표를 이해함- Python 개발 환경(Anaconda, VS Code)을 설치하고 구성함- Jupyter Notebook의 기본 사용법을 익힘 3) 학습내용: <ul style="list-style-type: none">- AI S/W 개발자 기초체력과정 개요 및 로드맵 소개- Python 3.x 및 Anaconda 설치, 가상환경 생성- VS Code 설치 및 Python 확장 프로그램 설정- Jupyter Notebook 실행 및 셀 실행 방법 실습	이론+실습	
2	1) 차시명: 공개 데이터셋 확보 및 데이터 생태계 이해 2) 학습목표: <ul style="list-style-type: none">- AI 학습에 필요한 공개 데이터셋의 종류와 특성을 파악함- AI 허브, Kaggle, UCI Repository 등 주요 데이터 플랫폼을 활용함- 데이터셋 다운로드 및 기본 구조 확인 방법을 습득함 3) 학습내용: <ul style="list-style-type: none">- 공공데이터포털, AI 허브 회원가입 및 API 키 발급- Kaggle 계정 생성 및 데이터셋 검색/다운로드- UCI ML Repository 데이터셋 구조 분석	이론+실습	

	- CSV, JSON, XML 등 데이터 형식의 이해		
3	1) 차시명: Python 기초 문법과 자료구조 2) 학습목표: <ul style="list-style-type: none"> - Python의 기본 문법과 자료형을 이해함 - 리스트, 딕셔너리, 튜플 등 자료구조를 활용함 - 조건문, 반복문, 함수 정의를 통한 프로그래밍 기초를 습득함 3) 학습내용: <ul style="list-style-type: none"> - 변수, 자료형(int, float, str, bool) 및 형변환 - 리스트, 튜플, 딕셔너리, 집합의 생성과 조작 - if-elif-else 조건문, for/while 반복문 - 함수 정의(def), 매개변수, 반환값 처리 	이론+실습	
4	1) 차시명: NumPy와Pandas 기초 2) 학습목표: <ul style="list-style-type: none"> - NumPy 배열의 생성과 연산 방법을 익힘 - Pandas DataFrame을 활용한 데이터 조작 기초를 습득함 - 데이터 읽기/쓰기 및 기본 탐색 방법을 이해함 3) 학습내용: <ul style="list-style-type: none"> - NumPy ndarray 생성, 인덱싱, 슬라이싱 - 배열 연산(브로드캐스팅, 유니버설 함수) - Pandas Series와DataFrame 구조 이해 - CSV/Excel 파일 읽기(read_csv, read_excel) 및head(), info(), describe() 활용 	이론+실습	실습과제 1: 공개 데이터셋 로드 및 기초 탐색
5	1) 차시명: 기술통계량의 기본과 의미 2) 학습목표: <ul style="list-style-type: none"> - 평균, 중앙값, 최빈값 등 중심경향치를 이해함 - 분산, 표준편차, 범위 등 산포도 지표를 계산함 - 기술통계량의 실무적 의미와 활용 방안을 파악함 3) 학습내용: <ul style="list-style-type: none"> - 중심경향치: mean, median, mode 계산 원리 - 산포도: variance, std, range, IQR 개념 - Pandas describe()와 수작업 계산 비교 실습 - 실제 공공데이터(서울시 대기질 데이터)를 활용한 기술통계 분석 	이론+실습	
6	1) 차시명: 데이터 시각화 기초- Matplotlib과Seaborn 2) 학습목표: <ul style="list-style-type: none"> - Matplotlib을 활용한 기본 차트 생성 방법을 익힘 - Seaborn을 통한 통계적 시각화 기법을 습득함 - 데이터 분포와 관계를 시각적으로 탐색함 3) 학습내용: <ul style="list-style-type: none"> - Matplotlib 기본 구조(figure, axes, plot) - 선 그래프, 막대 그래프, 히스토그램, 산점도 작성 - Seaborn을 활용한boxplot, violinplot, heatmap - 한글 폰트 설정 및 그래프 저장 방법 	이론+실습	
7	1) 차시명: 확률분포와 정규분포의 이해 2) 학습목표: <ul style="list-style-type: none"> - 확률의 기본 개념과 확률분포의 의미를 파악함 - 정규분포의 특성과 중심극한정리를 이해함 - Python으로 확률분포를 시뮬레이션함 3) 학습내용: <ul style="list-style-type: none"> - 확률의 정의, 조건부 확률, 베이즈 정리 개요 - 이산확률분포(이항분포, 포아송분포) 개념 - 연속확률분포와 정규분포(μ, σ)의 특성 - scipy.stats를 활용한 정규분포 시각화 및z-score 계산 	이론+실습	
8	1) 차시명: 가설검정의 기초- t-test 2) 학습목표: <ul style="list-style-type: none"> - 귀무가설과 대립가설의 개념을 이해함 - t-검정의 원리와 적용 조건을 파악함 - Python으로 일표본 및 독립표본t-검정을 수행함 3) 학습내용: <ul style="list-style-type: none"> - 가설검정의 논리: 귀무가설, 대립가설, 유의수준, p-value 	이론+실습	실습과제 2: 공공데이터 그룹 비교 분석

	<ul style="list-style-type: none"> - 일표본t-검정(one-sample t-test) 원리와 수식 - 독립표본t-검정(independent t-test) 적용 - <code>scipy.stats.ttest_1samp</code>, <code>ttest_ind</code>를 활용한 실습 		
9	1) 차시명: 분산분석(ANOVA)과 다중 비교 2) 학습목표: <ul style="list-style-type: none"> - 세 집단 이상의 평균 비교에ANOVA를 적용함 - F-통계량과 분산비의 의미를 이해함 - 사후검정(Post-hoc test)의 필요성과 방법을 습득함 3) 학습내용: <ul style="list-style-type: none"> - ANOVA의 기본 가정과F-분포 개념 - 일원분산분석(One-way ANOVA) 수행 - <code>scipy.stats.f_oneway</code>를 활용한 분석 실습 - Tukey HSD 사후검정을 통한 집단 간 차이 확인 	이론+실 습	
10	1) 차시명: 카이제곱검정과 범주형 데이터 분석 2) 학습목표: <ul style="list-style-type: none"> - 범주형 데이터의 특성과 분석 방법을 이해함 - 카이제곱 적합도 검정과 독립성 검정을 수행함 - 교차표(Contingency Table) 분석 기법을 습득함 3) 학습내용: <ul style="list-style-type: none"> - 범주형vs 연속형 데이터의 구분과 처리 - 카이제곱 적합도 검정(Goodness of fit) 원리 - 카이제곱 독립성 검정(Test of independence) - <code>pd.crosstab</code>과<code>scipy.stats.chi2_contingency</code> 실습 	이론+실 습	
11	1) 차시명: 상관분석과 인과관계 2) 학습목표: <ul style="list-style-type: none"> - 상관관계와 인과관계의 차이를 명확히 구분함 - 피어슨, 스피어만 상관계수를 계산하고 해석함 - 상관분석의 한계와 주의사항을 이해함 3) 학습내용: <ul style="list-style-type: none"> - 상관계수의 정의와 해석(-1 ~ +1 범위) - 피어슨 상관계수: 선형관계 측정 - 스피어만 순위 상관계수: 비선형 관계 포착 - <code>Pandas corr()</code>와<code>Seaborn heatmap</code>을 활용한 상관행렬 시각화 	이론+실 습	
12	1) 차시명: 데이터 전처리(I) - 결측치와 이상치 처리 2) 학습목표: <ul style="list-style-type: none"> - 결측치의 유형과 처리 전략을 파악함 - 이상치 탐지 방법(IQR, Z-score)을 이해함 - 실제 데이터셋에 결측치/이상치 처리를 적용함 3) 학습내용: <ul style="list-style-type: none"> - 결측치 확인: <code>isnull()</code>, <code>isna()</code>, <code>info()</code> 활용 - 결측치 처리: 삭제(<code>dropna</code>), 대체(<code>fillna</code>, 평균/중앙값/최빈값) - 이상치 탐지: IQR 방법, Z-score 기반 탐지 - AI 허브 공공데이터를 활용한 전처리 실습 	이론+실 습	실습과제 3: 결측치 /이상치 처리 보고 서 작성
13	1) 차시명: 데이터 전처리(II) - 피쳐 엔지니어링 2) 학습목표: <ul style="list-style-type: none"> - 범주형 변수 인코딩 기법을 습득함 - 수치형 변수 스케일링 방법을 이해함 - 파생 변수 생성을 통한 피쳐 엔지니어링을 수행함 3) 학습내용: <ul style="list-style-type: none"> - Label Encoding vs One-Hot Encoding 비교 - <code>StandardScaler</code>, <code>MinMaxScaler</code>, <code>RobustScaler</code> 적용 - 날짜/시간 데이터에서 파생 변수 추출 - <code>scikit-learn</code>의<code>preprocessing</code> 모듈 활용 실습 	이론+실 습	
14	1) 차시명: 시계열 데이터 전처리와 분석 기초 2) 학습목표: <ul style="list-style-type: none"> - 시계열 데이터의 특성과 구성 요소를 이해함 - 날짜/시간 인덱싱과 리샘플링 기법을 익힘 - 이동평균과 추세 분석 방법을 습득함 3) 학습내용:	이론+실 습	

	<ul style="list-style-type: none"> - 시계열 데이터의 구성: 추세, 계절성, 잔차 - Pandas datetime 인덱스 설정 및 to_datetime() - resample()을 활용한 시간 단위 집계 - rolling()을 활용한 이동평균선 계산 및 시각화 		
15	<p>1) 차시명: 텍스트 데이터 전처리 기초</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 텍스트 데이터의 특성과 전처리 필요성을 이해함 - 토큰화, 정규화, 불용어 제거 등 기본 기법을 습득함 - 한국어 텍스트 처리의 특수성을 파악함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 텍스트 정규화: 소문자 변환, 특수문자 제거 - 토큰화(Tokenization): 단어/문장 단위 분리 - 불용어(Stopwords) 제거와 어간 추출(Stemming) - KoNLPy를 활용한 한국어 형태소 분석 맛보기 	이론+실습	
16	<p>1) 차시명: 분류(Classification)와 회귀(Regression)의 이해</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 지도학습의 두 가지 유형을 명확히 구분함 - 분류 문제와 회귀 문제의 특성을 이해함 - 적절한 문제 유형 선택 기준을 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 지도학습 개요: 입력(X)과 레이블(y)의 관계 - 분류: 이진 분류 vs 다중 분류, 대표 사례 - 회귀: 연속값 예측, 선형 회귀 개념 - 문제 유형에 따른 평가지표 차이(정확도 vs RMSE) 	이론	
17	<p>1) 차시명: 선형회귀와 다중선형회귀</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 선형회귀의 수학적 원리를 이해함 - 최소제곱법(OLS)의 개념과 계산 과정을 습득함 - 다중선형회귀 모델을 구현하고 해석함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 단순선형회귀: $y = ax + b$, 기울기와 절편의 의미 - 최소제곱법 원리와 잔차(residual) 개념 - 다중선형회귀: 여러 독립변수를 활용한 예측 - scikit-learn LinearRegression 구현 및 계수 해석 	이론+실습	
18	<p>1) 차시명: 로지스틱 회귀와 분류 모델 평가</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 로지스틱 회귀의 원리와 시그모이드 함수를 이해함 - 혼동행렬(Confusion Matrix) 기반 평가지표를 습득함 - 정밀도, 재현율, F1-score의 의미를 파악함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 로지스틱 회귀: 확률 출력과 시그모이드 함수 - 결정 경계(Decision Boundary)의 개념 - 혼동행렬: TP, TN, FP, FN 이해 - Accuracy, Precision, Recall, F1-score 계산 및 해석 	이론+실습	실습과제 4. 이진분류 모델 구축 및 평가
19	<p>1) 차시명: 의사결정나무와 앙상블 기법</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 의사결정나무의 분할 기준과 구조를 이해함 - 과적합 방지를 위한 가지치기 기법을 습득함 - 랜덤포레스트의 원리와 장점을 파악함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 의사결정나무: 지니 불순도, 정보 이득 기준 - 트리 시각화와 특성 중요도(feature importance) - 과적합 문제와 가지치기(pruning) 전략 - 랜덤포레스트: 배깅(Bagging)과 앙상블 개념 	이론+실습	
20	<p>1) 차시명: SVM과 KNN, 그리고 모델 선택</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - SVM(Support Vector Machine)의 기본 원리를 이해함 - KNN(K-Nearest Neighbors)의 작동 방식을 습득함 - 교차검증과 하이퍼파라미터 튜닝 기법을 익힘 <p>3) 학습내용:</p>	이론+실습	

	<ul style="list-style-type: none"> - SVM: 마진 최대화, 커널 트릭 개념 - KNN: 거리 기반 분류, K값 선택의 중요성 - train_test_split과 교차검증(Cross-Validation) - GridSearchCV를 활용한 하이퍼파라미터 최적화 		
21	<p>1) 차시명: 인공지능과 머신러닝, 딥러닝의 관계</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - AI, 머신러닝, 딥러닝의 개념적 위계를 이해함 - 지도학습, 비지도학습, 강화학습의 차이를 파악함 - 딥러닝이 등장한 배경과 핵심 특징을 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 인공지능의 역사와 발전 단계 - AI > ML > DL 관계도와 각 영역의 특성 - 지도/비지도/강화학습 분류 체계 - 딥러닝의 등장 배경: 빅데이터, GPU, 알고리즘 혁신 	이론	
22	<p>1) 차시명: 신경망의 기초- 퍼셉트론과 활성화 함수</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 퍼셉트론의 구조와 동작 원리를 이해함 - 다양한 활성화 함수의 특성과 역할을 파악함 - 단층 퍼셉트론의 한계와 다층 퍼셉트론의 필요성을 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 퍼셉트론: 입력, 가중치, 편향, 출력 구조 - 활성화 함수: Sigmoid, ReLU, Softmax 비교 - XOR 문제와 단층 퍼셉트론의 한계 - 다층 퍼셉트론(MLP) 아키텍처 개요 	이론+실습	
23	<p>1) 차시명: 손실함수와 경사하강법</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 손실함수의 개념과 종류를 이해함 - 경사하강법의 수학적 원리를 파악함 - 학습률(Learning Rate)의 중요성을 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 손실함수: MSE, Cross-Entropy 개념과 수식 - 경사하강법: 미분을 통한 최적화 원리 - 학습률 선택에 따른 학습 양상 변화 - NumPy를 활용한 간단한 경사하강법 구현 	이론+실습	
24	<p>1) 차시명: 역전파 알고리즘의 이해</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 역전파(Backpropagation)의 동작 원리를 이해함 - 연쇄법칙(Chain Rule)을 활용한 기울기 계산을 습득함 - 가중치 업데이트 과정을 코드로 구현함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 순전파(Forward Propagation) 과정 복습 - 역전파: 출력층에서 입력층으로의 오차 전파 - 연쇄법칙을 활용한 편미분 계산 - 간단한 2층 신경망의 역전파 수동 계산 실습 	이론+실습	실습과제 5: 역전파 알고리즘 직접 구현
25	<p>1) 차시명: 기울기 소실 문제와 해결 방안</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 기울기 소실/폭발 문제의 원인을 이해함 - ReLU와 He 초기화의 역할을 파악함 - 배치 정규화(Batch Normalization)의 효과를 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - Sigmoid의 포화 영역과 기울기 소실 현상 - ReLU 활성화 함수의 장점과 Dead ReLU 문제 - 가중치 초기화: Xavier, He 초기화 기법 - 배치 정규화의 원리와 적용 효과 	이론+실습	
26	<p>1) 차시명: TensorFlow/Keras 기초와 첫 번째 신경망</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - TensorFlow/Keras의 기본 구조를 이해함 - Sequential 모델을 활용한 신경망 구축을 습득함 - MNIST 손글씨 분류 모델을 직접 구현함 <p>3) 학습내용:</p>	이론+실습	

	<ul style="list-style-type: none"> - TensorFlow/Keras 설치 및 환경 설정 - Sequential API를 활용한 모델 정의 - compile(), fit(), evaluate() 메서드 활용 - MNIST 데이터셋을 활용한 손글씨 분류 실습 		
27	<p>1) 차시명: CNN 기초- 합성곱과 풀링</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 합성곱 연산의 원리와 특징 추출 개념을 이해함 - 풀링(Pooling)의 역할과 종류를 파악함 - CNN 아키텍처의 기본 구조를 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 합성곱(Convolution) 연산: 필터/커널 개념 - 스트라이드(Stride)와 패딩(Padding) 설정 - 맥스 풀링(Max Pooling)과 평균 풀링 - Convolution Layer - Pooling Layer - FC Layer 구조 	이론+실습	
28	<p>1) 차시명: CNN 실습- 이미지 분류 모델 구축</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - Keras를 활용한CNN 모델을 구현함 - CIFAR-10 데이터셋을 활용한 이미지 분류를 수행함 - 데이터 증강(Data Augmentation) 기법을 적용함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - Keras Conv2D, MaxPooling2D 레이어 활용 - CIFAR-10 데이터셋 로드 및 전처리 - 모델 훈련과 학습 곡선 분석 - ImageDataGenerator를 활용한 데이터 증강 	이론+실습	실습과제 6: CNN 이미지 분류 프로젝트
29	<p>1) 차시명: 전이학습과 사전훈련 모델 활용</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 전이학습(Transfer Learning)의 개념과 장점을 이해함 - VGG, ResNet 등 사전훈련 모델의 특징을 파악함 - Fine-tuning 기법을 활용한 모델 재훈련을 수행함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 전이학습의 원리: 사전학습된 가중치 활용 - ImageNet 기반 사전훈련 모델 소개 - Keras applications를 활용한 모델 로드 - 특성 추출(Feature Extraction) vs Fine-tuning 비교 	이론+실습	
30	<p>1) 차시명: 객체 탐지 개요- YOLO 아키텍처</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 객체 탐지(Object Detection)의 개념을 이해함 - YOLO 알고리즘의 동작 원리를 파악함 - 사전훈련된 YOLO 모델을 활용한 탐지를 수행함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 객체 탐지vs 이미지 분류: 바운딩 박스 개념 - YOLO: You Only Look Once 원리 - 앵커 박스(Anchor Box)와 비최대 억제(NMS) - YOLOv5/v8을 활용한 객체 탐지 데모 	이론+실습	
31	<p>1) 차시명: RNN 기초- 순환 신경망의 이해</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - 시퀀스 데이터 처리를 위한RNN의 구조를 이해함 - RNN의 순환 구조와 은닉 상태(Hidden State) 개념을 파악함 - RNN의 장단점과 한계를 습득함 <p>3) 학습내용:</p> <ul style="list-style-type: none"> - 시퀀스 데이터의 특성과 처리 필요성 - RNN 셀 구조: 입력, 은닉 상태, 출력 - 시간축 전개(Unfolding)와 역전파(BPTT) - 장기 의존성 문제와 기울기 소실/폭발 	이론+실습	
32	<p>1) 차시명: LSTM과 시계열 예측</p> <p>2) 학습목표:</p> <ul style="list-style-type: none"> - LSTM(Long Short-Term Memory)의 게이트 구조를 이해함 - LSTM을 활용한 시계열 데이터 예측 모델을 구축함 - 주가/기온 예측 등 실제 사례에 적용함 	이론+실습	실습과제 7: LSTM 시계열 예측 프로젝트

	3) 학습내용: <ul style="list-style-type: none"> - LSTM 셀: 망각 게이트, 입력 게이트, 출력 게이트 - Keras LSTM 레이어 활용법 - 시계열 데이터 윈도우 생성 및 전처리 - 실제 기상 데이터를 활용한 기온 예측 실습 		
33	1) 차시명: 비지도학습(I) - 군집화(Clustering) 2) 학습목표: <ul style="list-style-type: none"> - 비지도학습의 개념과 활용 분야를 이해함 - K-Means 클러스터링 알고리즘을 구현함 - 실루엣 점수를 통한 군집 품질 평가를 수행함 3) 학습내용: <ul style="list-style-type: none"> - 비지도학습 개요: 레이블 없는 데이터 학습 - K-Means 알고리즘: 중심점 갱신 과정 - 엘보우 기법(Elbow Method)을 활용한 K 선택 - 실루엣 분석(Silhouette Analysis)과 시각화 	이론+실습	
34	1) 차시명: 비지도학습(II) - 차원축소(PCA) 2) 학습목표: <ul style="list-style-type: none"> - 차원의 저주(Curse of Dimensionality) 개념을 이해함 - PCA(주성분분석)의 수학적 원리를 파악함 - 고차원 데이터 시각화에 PCA를 적용함 3) 학습내용: <ul style="list-style-type: none"> - 차원축소의 필요성과 장점 - 공분산 행렬과 고유값 분해 개념 - scikit-learn PCA를 활용한 차원축소 실습 - 2D/3D 시각화를 통한 데이터 구조 탐색 	이론+실습	
35	1) 차시명: GAN 기초- 생성적 적대 신경망 2) 학습목표: <ul style="list-style-type: none"> - GAN(Generative Adversarial Network)의 원리를 이해함 - 생성자(Generator)와 판별자(Discriminator)의 역할을 파악함 - 간단한 GAN 모델을 구현하고 이미지를 생성함 3) 학습내용: <ul style="list-style-type: none"> - GAN의 기본 아이디어: 생성자 vs 판별자 경쟁 - 손실함수 설계와 학습 과정 - Keras를 활용한 간단한 GAN 구현 - MNIST 손글씨 이미지 생성 실습 	이론+실습	
36	1) 차시명: Transformer와 LLM의 이해 2) 학습목표: <ul style="list-style-type: none"> - Transformer 아키텍처의 핵심 구조를 이해함 - Self-Attention 메커니즘의 동작 원리를 파악함 - GPT, BERT 등 대규모 언어모델의 기반을 습득함 3) 학습내용: <ul style="list-style-type: none"> - Attention 메커니즘: Query, Key, Value 개념 - Self-Attention과 Multi-Head Attention - Positional Encoding의 역할 - GPT, BERT, LLM의 발전 흐름과 응용 분야 	이론+실습	실습과제 8: Hugging Face Transformers 활용
37	1) 차시명: Streamlit을 활용한 AI 웹 애플리케이션 2) 학습목표: <ul style="list-style-type: none"> - Streamlit의 기본 구조와 컴포넌트를 이해함 - 학습된 모델을 웹 인터페이스로 배포하는 방법을 습득함 - 사용자 입력을 받아 예측 결과를 표시함 3) 학습내용: <ul style="list-style-type: none"> - Streamlit 설치 및 기본 구조 이해 - 텍스트, 차트, 입력 위젯 컴포넌트 활용 - pickle/joblib을 활용한 모델 저장 및 로드 - 이미지 분류 모델 웹앱 구축 실습 	이론+실습	
38	1) 차시명: FastAPI를 활용한 AI 모델 API 구축 2) 학습목표: <ul style="list-style-type: none"> - REST API의 기본 개념을 이해함 - FastAPI를 활용한 예측 API 서버를 구축함 	이론+실습	

	- Swagger UI를 통한API 테스트 방법을 습득함 3) 학습내용: - REST API 개념: GET, POST 메서드 - FastAPI 설치 및 기본 라우팅 설정 - Pydantic을 활용한 요청/응답 스키마 정의 - AI 모델 추론API 엔드포인트 구현		
39	1) 차시명: AI 프로젝트 종합 실습(I) 2) 학습목표: - 전 과정에서 학습한 내용을 종합적으로 적용함 - 공개 데이터셋을 활용한End-to-End 프로젝트를 수행함 - 문제 정의부터 모델 배포까지 전체 파이프라인을 구축함 3) 학습내용: - 프로젝트 주제 선정 및 문제 정의 - 데이터 수집, 탐색적 데이터 분석(EDA) - 데이터 전처리 및 피처 엔지니어링 - 모델 선택, 학습 및 하이퍼파라미터 튜닝	실습	종합 프로젝트
40	1) 차시명: AI 프로젝트 종합 실습(II) 및 과정 마무리 2) 학습목표: - 구축된 모델을 웹 서비스로 배포함 - 프로젝트 결과를 문서화하고 발표 자료를 작성함 - AI 기초체력과정의 학습 내용을 정리하고 향후 학습 방향을 설정함 3) 학습내용: - Streamlit/FastAPI를 활용한 모델 배포 - 프로젝트 보고서 작성 및 발표 준비 - 학습 내용 총정리 및 핵심 개념 복습 - AI S/W 개발자 심화 과정 안내 및 학습 로드맵	실습	최종 프로젝트 제출

< 평가 문항>

◆ 레포트 주제(1주~ 20차시 내용)

공공데이터포털 또는AI 허브에서 관심 분야의 데이터셋을 선정하여, 데이터 수집→ 탐색적 데이터 분석(EDA) → 전처리→ 머신러닝 모델 적용→ 결과 해석의 전 과정을 수행하고, 각 단계별 방법론과 결과를 보고서로 정리함.

◆ 1차 토론(5주~ 10차시 내용)

주제: "데이터 분석에서p-value의 한계와 효과크기(Effect Size)의 중요성"에 대해 논의함. 통계적 유의성만으로 결론을 내리는 것의 문제점과 실무에서의 대안을 탐색함.

◆ 2차 토론(21주~ 36차시 내용)

주제: "전이학습(Transfer Learning)이AI 민주화에 기여하는가?"에 대해 논의함. 사전훈련 모델 활용의 장점과 잠재적 문제점(편향, 블랙박스 등)을 다각도로 분석함.

◆ 1차 퀴즈(OX, 1주~ 20차시 내용)

문제: "정규분포에서 평균으로부터±1 표준편차 범위 내에 약68%의 데이터가 포함된다."
정답: O

◆ 2차 퀴즈(OX, 21주~ 40차시 내용)

문제: "LSTM은 기울기 소실 문제를 해결하기 위해 게이트 메커니즘을 도입한 순환 신경망 구조이다."
정답: O