

인공지능 프로젝트과정 이해와 산출물의 요소



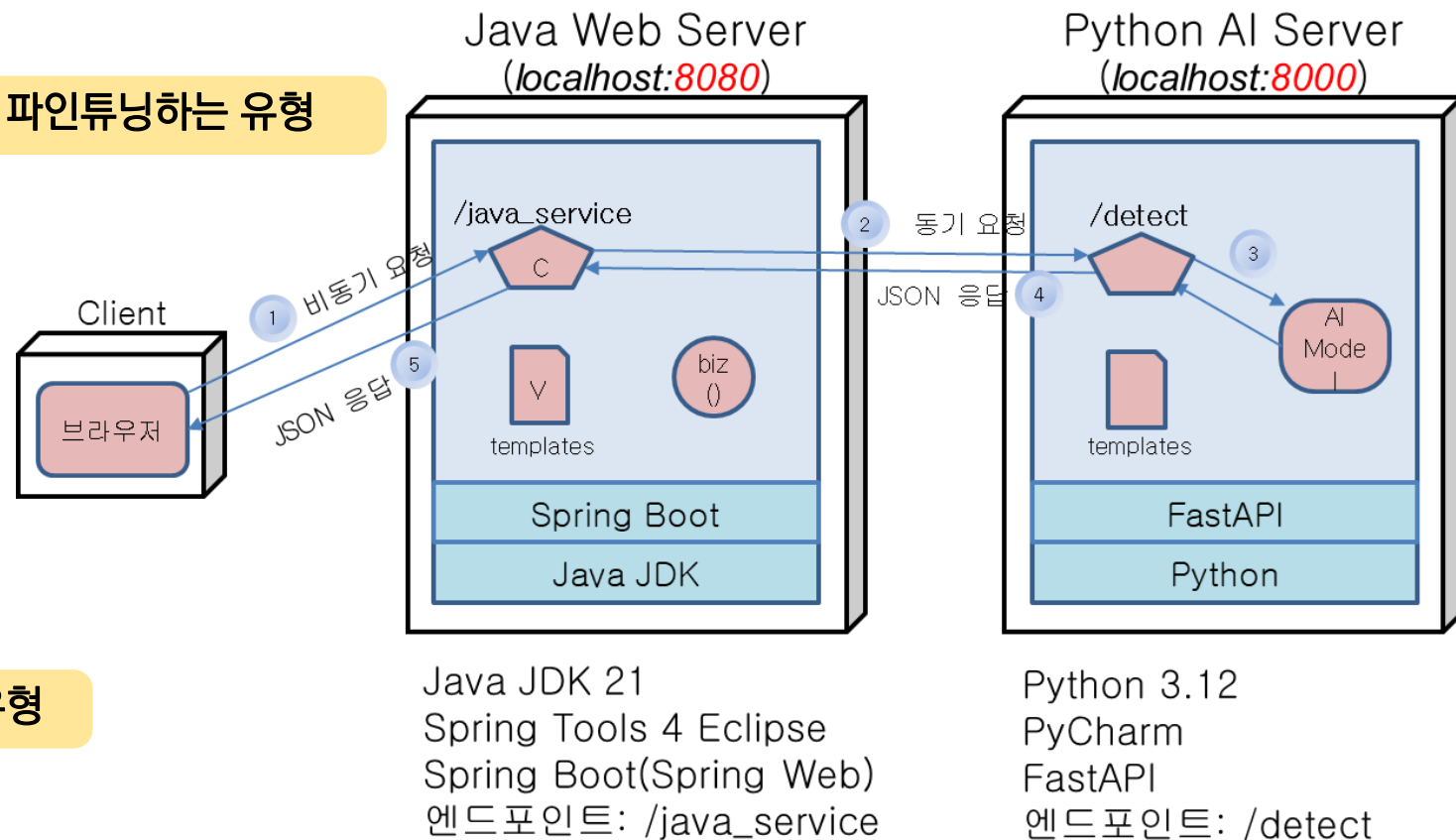
AI PROJECT
DOCUMENTS



인공지능 프로젝트 유형

인공지능 프로젝트과정 이해와 산출물의 요소

- ✓ 모델을 직접 구현하고 기존 모델과 비교하는 유형
- ✓ 이미 구현된 모델을 인공지능 플랫폼에서 가져와 파인튜닝하는 유형
- ✓ OpenAPI를 사용하는 유형
- ✓ 웹으로 인공지능 서비스를 제공하는 유형
- ✓ 자바 웹과 파이썬 인공지능 플랫폼을 연동하는 유형



주제 선정

인공지능 프로젝트과정 이해와 산출물의 요소



프로젝트 주제

- ▶ 프로젝트명
- ▶ 프로젝트 개요
- ▶ 팀 이름 및 팀원 구성
- ▶ 데이터 출처
- ▶ 사용 기술 및 아키텍처



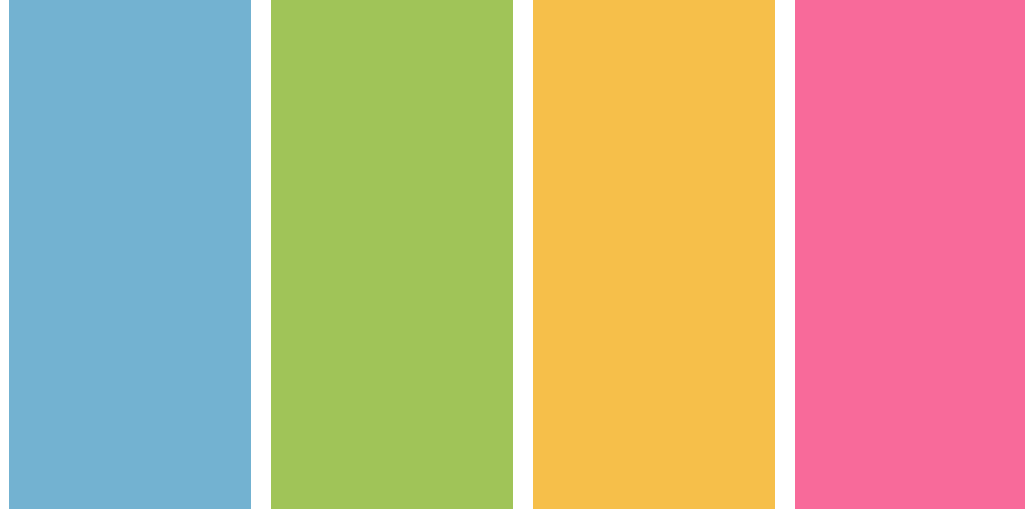
- 프로젝트 주제가 정해졌으면
 - 어떤 모델을 이용할지(찾아서 사용하거나, 아키텍처를 정의하거나) 구상해야 하고,
 - 프로젝트 개요를 작성하셔야 합니다.
 - 프로젝트 개요에는 아래 내용이 들어가도록 작성하세요.
 1. 주제 및 선정 배경(또는 기획의도)
 2. 프로젝트에 사용한 개발환경(장비 및 재료)과 기술 스택을 기록하세요. (개발환경은 버전이 기록되어야 합니다.)
 3. 프로젝트 구조를 구상했다면 도식화하세요.

페이지 번호를 삽입하세요.
(위치는 어디가 되든지 상관 없습니다.)

작성 방법

인공지능 프로젝트과정 이해와 산출물의 요소

- ▶ 본 포트폴리오 양식은 프로젝트 팀 별로 각각 작성하여 제출
- ▶ 프로젝트 수행 과정 및 결과에 대해서는 제공된 목차 및 세부 항목별 작성요령을 참조하여 작성
- ▶ 프로젝트 특성에 따라 기본적인 구성을 유지한 상태에서 제공 양식을 보완하거나 추가하여 작성할 수 있음
- ▶ 본 슬라이드의 템플릿을 사용하는 것은 지양
- ▶ 작성 예시 및 작성요령 등은 모두 삭제 후 제출



인공지능 프로젝트 발표 슬라이드 예시

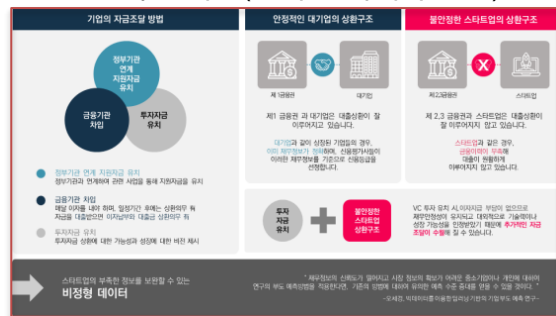
인공지능 프로젝트과정 이해와 산출물의 요소



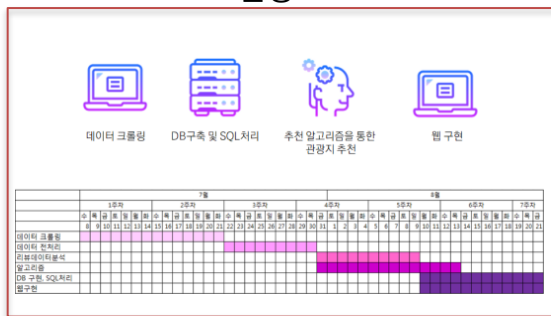
프로젝트 슬라이드 예시

인공지능 프로젝트과정 이해와 산출물의 요소

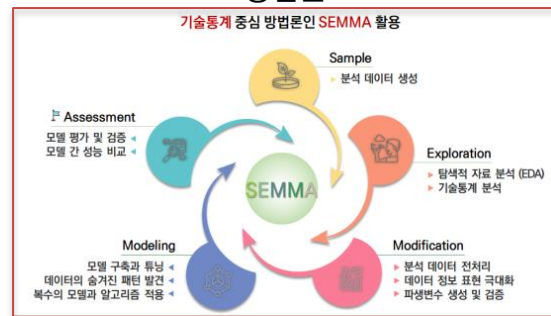
프로젝트 개요(도메인 이해가 필요)



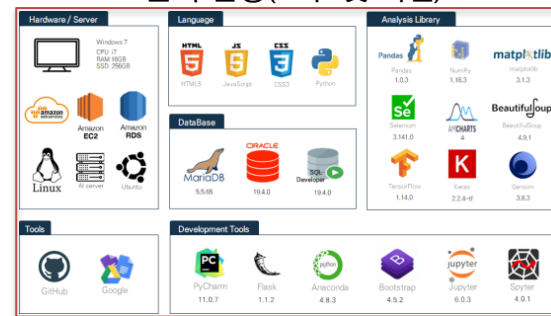
일정표



방법론



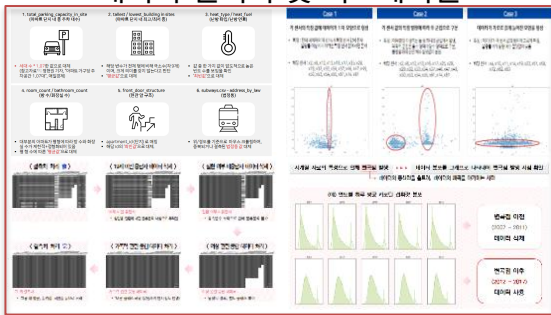
분석 환경(도구 및 버전)



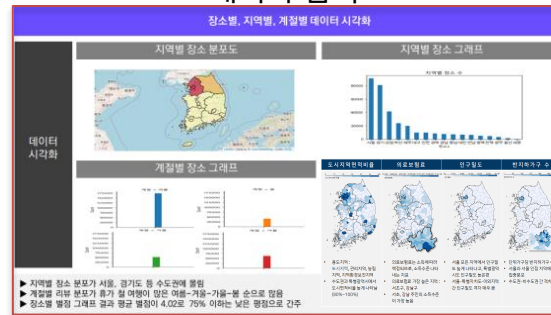
데이터 출처



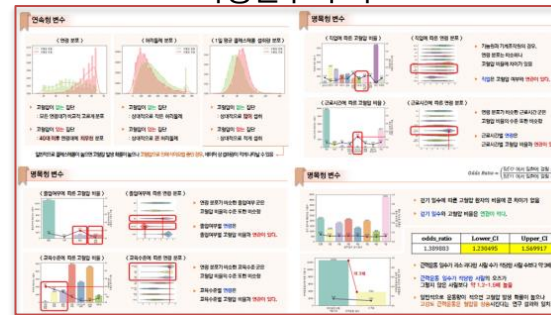
데이터 전처리 및 어노테이션



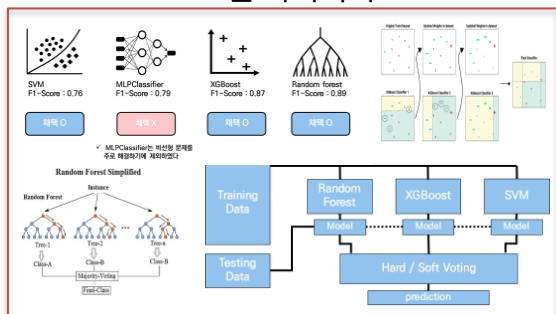
데이터 탐색



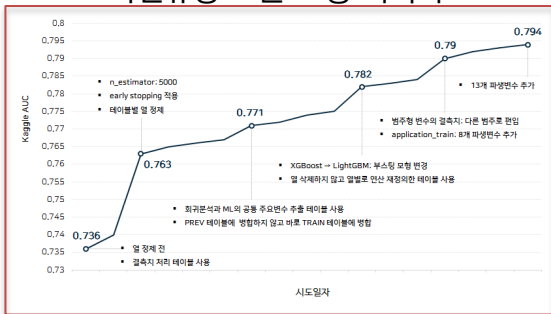
파생변수 추가



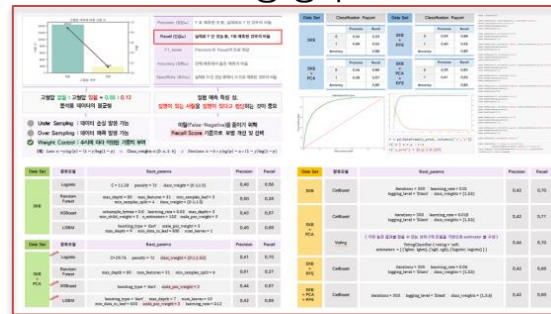
모델 아키텍처



파인튜닝 또는 모형 최적화



모형 평가



모형별 평가 비교표

모형명	Baseline	Library	Validation set	Without flow, action	Total flow Augmentation (train)	Total flow Augmentation (test)	Ensemble	Integration + Augmentation	Ensemble	Integration + Augmentation	Score
#1	CNN	ResNet	X	X	X	X	X	X	X	X	0.277
#2	ResNet	ResNet	X	X	X	X	X	X	X	X	0.290
#3	ResNet	ResNet	X	X	X	X	X	X	X	X	0.565
#4	ResNet	ResNet	X	X	X	X	X	X	X	X	0.562
#5	ResNet	ResNet	X	X	X	X	X	X	X	X	0.588
#6	ResNet	ResNet	X	X	X	X	X	X	X	X	0.562
#7	ResNet	ResNet	X	X	X	X	X	X	X	X	0.582
#8	ResNet	ResNet	X	X	X	X	X	X	X	X	0.577
#9	ResNet	ResNet	X	X	X	X	X	X	X	X	0.614
#10	ResNet	ResNet	X	X	X	X	X	X	X	X	0.590
#11	ResNet	ResNet	X	X	X	X	X	X	X	X	0.588
#12	ResNet	ResNet	X	X	X	X	X	X	X	X	0.562
#13	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#14	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#15	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#16	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#17	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#18	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#19	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#20	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#21	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#22	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#23	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#24	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#25	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#26	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#27	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#28	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#29	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#30	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#31	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#32	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#33	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#34	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#35	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#36	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#37	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#38	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#39	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#40	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#41	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#42	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#43	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#44	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#45	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#46	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#47	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#48	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#49	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#50	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#51	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#52	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#53	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#54	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#55	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#56	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#57	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#58	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#59	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#60	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#61	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#62	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#63	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#64	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#65	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#66	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#67	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#68	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#69	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#70	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#71	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#72	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#73	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#74	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#75	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#76	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#77	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#78	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#79	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#80	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#81	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#82	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#83	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#84	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#85	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#86	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#87	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#88	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#89	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#90	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#91	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#92	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#93	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#94	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#95	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#96	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#97	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#98	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#99	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573
#100	ResNet	ResNet	X	X	X	X	X	X	X	X	0.573



- 프로젝트 주제(프로젝트 명)
- 팀명
- 소속 및 이름

다양한 방법으로 작성 가능함

PPT의 크기는 Wide(권장) 또는 A4가로

주제(프로젝트명)

소속 및 이름

※ 양식은 예시로 자유롭게 변경 가능하나, 목차 안에 구성된 내용은 포함되도록 작성(해당 템플릿 활용 지양)



프로젝트 흐름도 또는 시스템 아키텍처

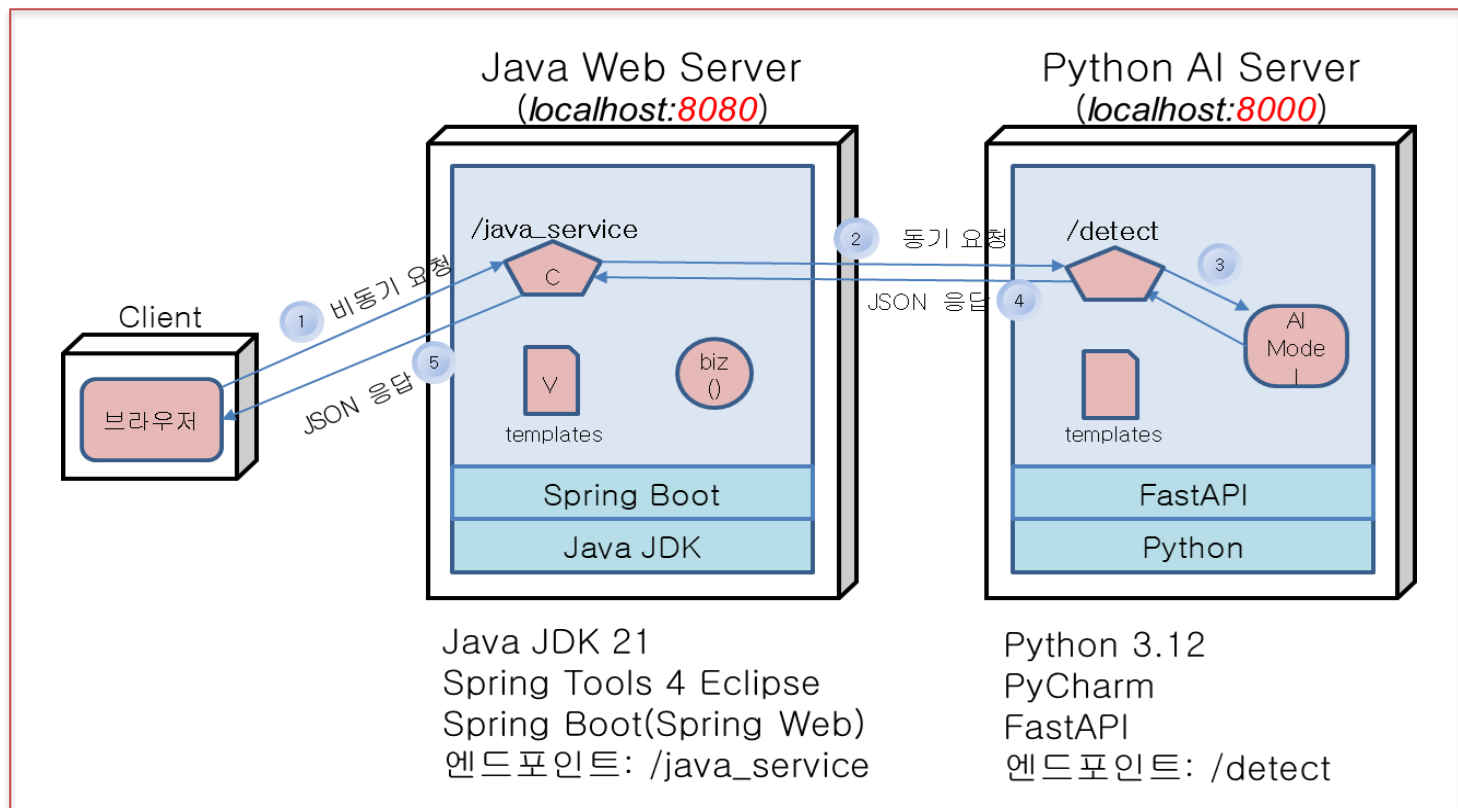
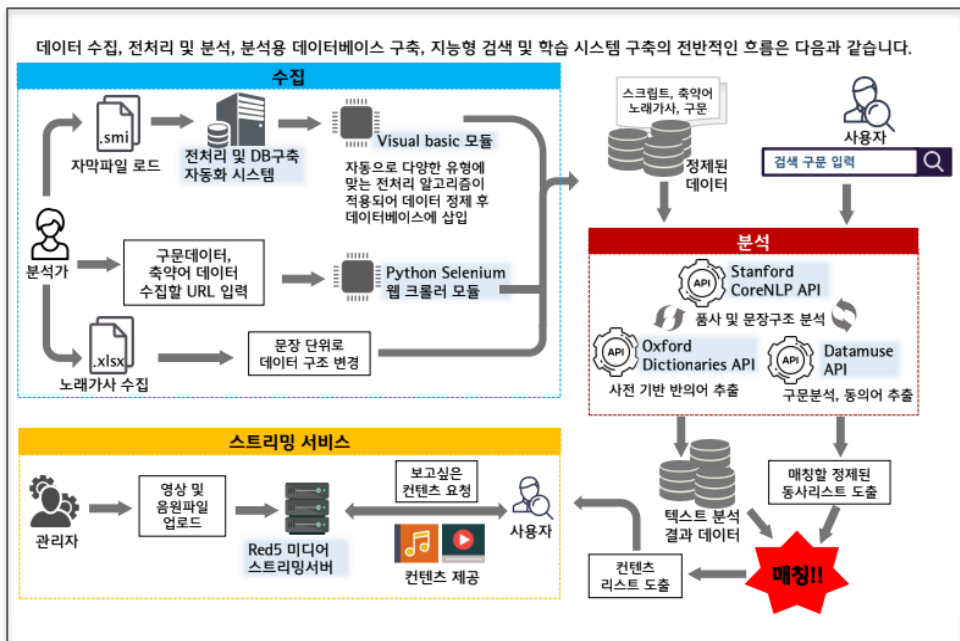
인공지능 프로젝트과정 이해와 산출물의 요소

- 프로젝트의 전체 흐름을 한눈에 파악할 수 있도록 프로젝트 흐름도를 작성
- 직관적이고 자세하기 작성
- 분석의 절차 및 시스템의 구성도를 포함할 수 있음



무엇보다 중요한 것은 전체 프로젝트 흐름도입니다.

다양한 방법으로 작성 가능함



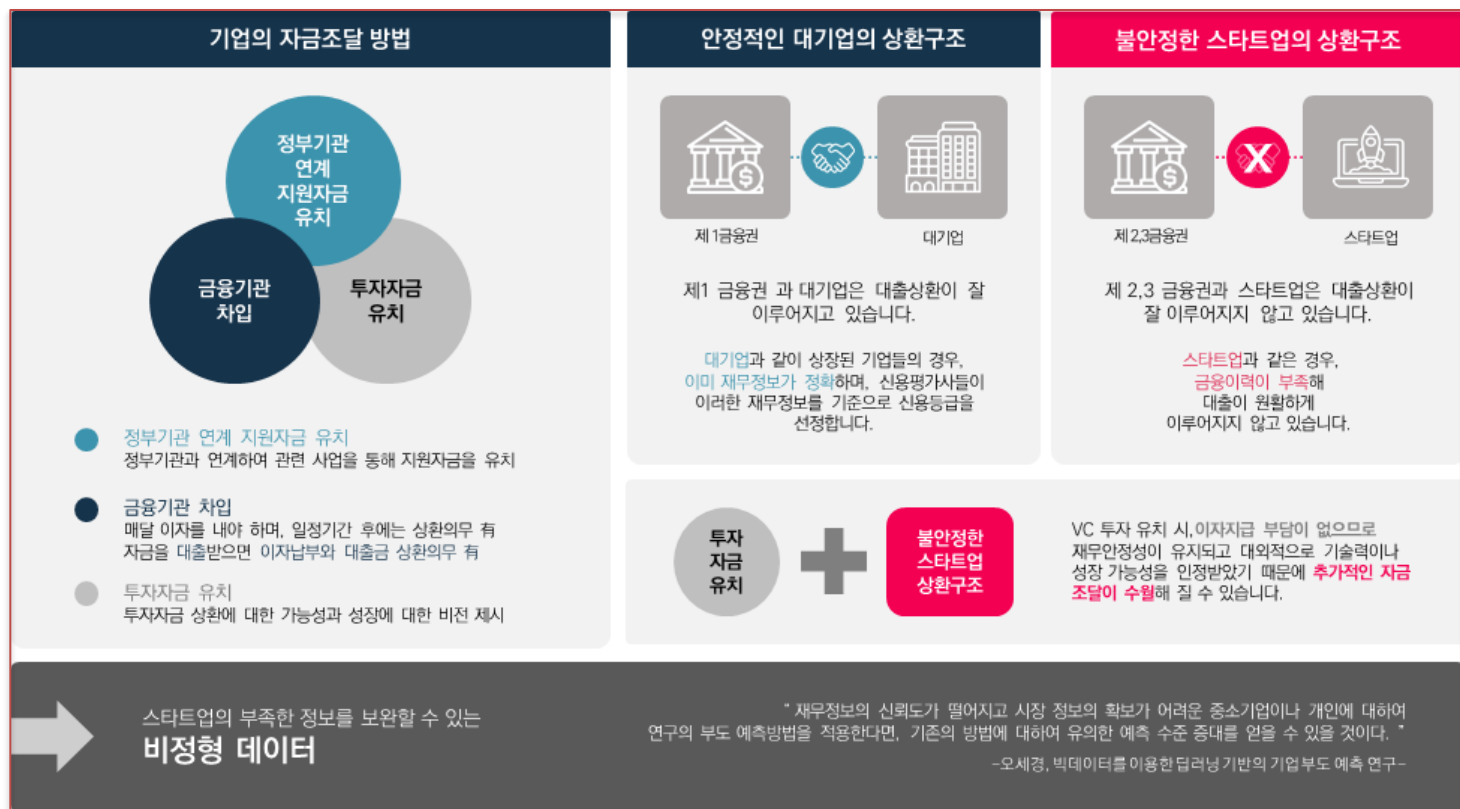
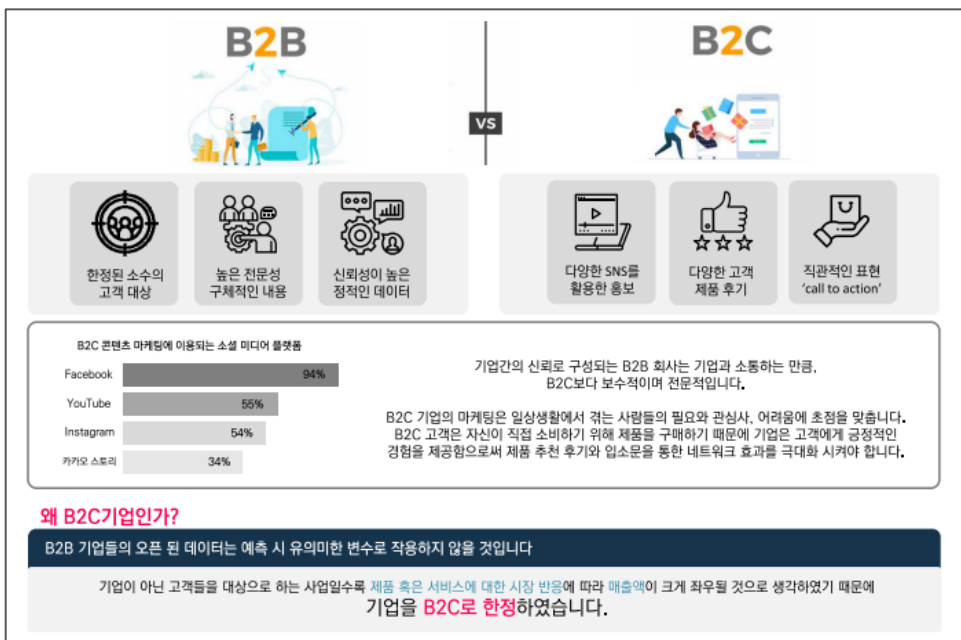


프로젝트 개요

인공지능 프로젝트과정 이해와 산출물의 요소

- 프로젝트의 목적 및 개요에 대해 기술
- 프로젝트 주제 및 선정 배경(기획의도 등) 및 기대효과
- 프로젝트 개요(프로젝트 구현 내용, 컨셉, 훈련내용과의 관련성 등)
- 활용 장비 및 재료(개발 환경 등)
- 프로젝트 구조

B2B, B2C, As-is To-be 등이 다음 슬라이드에 포함될 수 있음





프로젝트 팀 구성 및 역할

인공지능 프로젝트과정 이해와 산출물의 요소

- [프로젝트 팀 구성 및 역할]은 프로젝트를 기본 단위로 작성하며 팀원의 수에 따라 칸을 추가/삭제할 수 있다.
- 담당 업무 : 팀원 별로 해당 프로젝트를 진행하면서 주도적으로 참여한 부분을 중심으로 작성

팀원이 1명일 경우도 본인의 역할을 모두 기록

훈련생	역할	담당 업무
홍길동	팀장 및 팀원	▶ 데이터 정제 및 정규화
		▶ 모바일 서비스 테스트
		▶ 모바일 플랫폼 구현
		▶ 외부 데이터 수집
		▶ 서비스 시스템 설계
		▶ 텍스트 마이닝

훈련생	역할	담당 업무
김○○	팀장	▶ 데이터 정제 및 정규화 ▶ 모바일 서비스 테스트
박○○	팀원	▶ 모바일 플랫폼 구현 ▶ 외부 데이터 수집
정○○	팀원	▶ 서비스 시스템 설계 ▶ 텍스트 마이닝



프로젝트 수행 절차 및 방법

인공지능 프로젝트과정 이해와 산출물의 요소

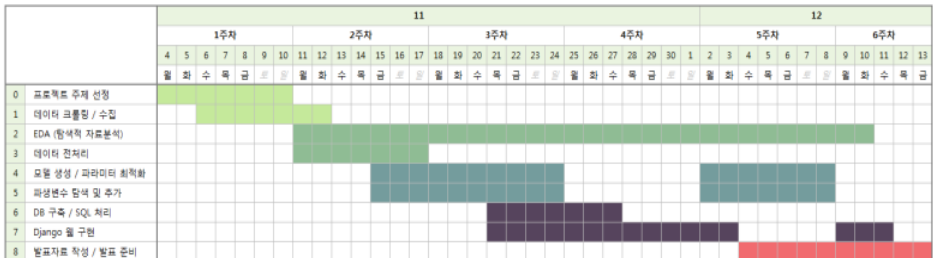
- [프로젝트 수행 절차 및 방법]은 프로젝트의 사전 기획과 프로젝트 수행 및 완료 과정으로 나누어서 작성.
- 프로젝트 수행 절차를 도식화하여 제시하거나, 더 효과적으로 전달하는 방법 등이 있다면 수정하여 작성 가능
- 기획 단계에서 도출된 주제와 아이디어를 기반으로 실제 프로젝트를 수행한 세부적인 기간과 활동 내용 작성

프로젝트 목적/목표와 함께 한 장의 장표로 가능

프로젝트 목적 : 교육 내용 전반을 활용한 예측 모델 개발 및 웹 서비스 프로토타입 구현

프로젝트 목표 : 섭취 식품 및 건강 설문 응답 데이터에서 질환 보유자의 패턴을 학습하고, 이를 근거로 일반 사용자의 질환 가능성 예측 결과를 제공할 수 있다.

프로젝트 범위 : 데이터 크롤링 / DB구축 & SQL처리 / 머신러닝 & 딥러닝 / 질환 예측 / 웹 구현



구분	기간	활동	비고
사전 기획	▶ O/O(월)~O/O(금)	▶ 프로젝트 기획 및 주제 선정 ▶ 기획안 작성	▶ 아이디어 선정
데이터 수집	▶ O/O(월)~O/O(금)	▶ 필요 데이터 및 수집 절차정의 ▶ 외부 데이터 수집	▶ OO 데이터 협조
데이터 전처리	▶ O/O(월)~O/O(금)	▶ 데이터 정제 및 정규화	
모델링	▶ O/O(월)~O/O(금)	▶ 모형 구현	
서비스 구축	▶ O/O(월)~O/O(금)	▶ 모바일 서비스 시스템 설계 ▶ 모바일 플랫폼 구현	▶ 최적화, 오류 수정
총 개발기간	▶ O/O(월)~O/O(금)(총 7주)	-	-



방법론

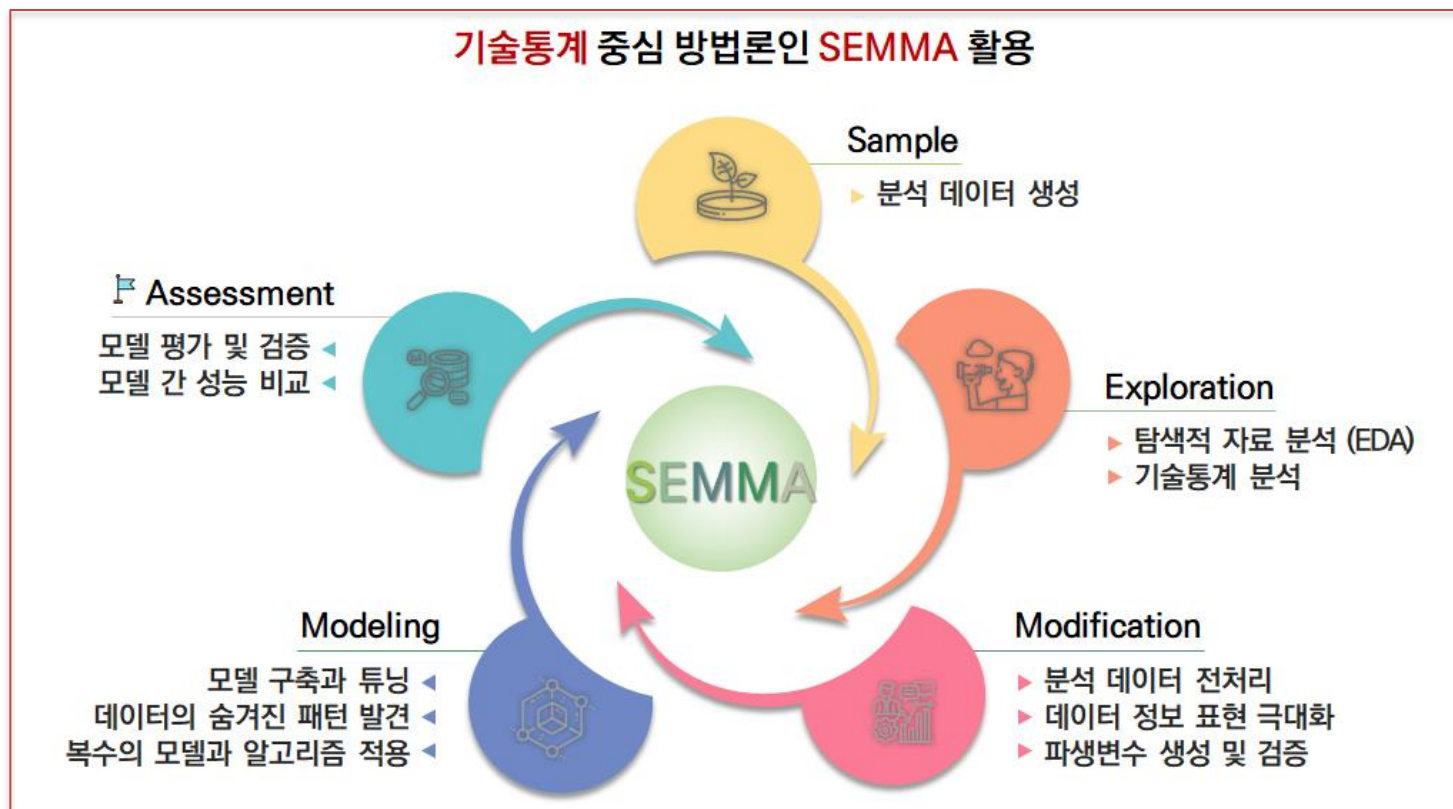
인공지능 프로젝트과정 이해와 산출물의 요소

- 프로젝트 방법론을 정의 함(선택사항)
- SEMMA 방법론만 있는 것은 아님
- 데이터 이해 단계가 포함되어야 한다면 CRISP-DM 방법론 사용

CRISP-DM
































기술통계 중심 방법론인 SEMMA 활용





- 분석 환경 또는 개발환경에 사용한 도구들 정리
- 이름과 버전을 명시해야 함
- 시스템 및 H/W
- 프로그래밍 언어 및 데이터베이스
- 개발도구, 라이브러리, 프레임워크

하드웨어, 서버, 데이터베이스 등	사용 언어	분석 라이브러리
 <p>CPU i7 RAM 16GB SSD 256GB</p>  <p>WebServer (Python/Fask)</p>  <p>Oracle Data Base</p>	 <p>HTML</p>  <p>JavaScript</p>  <p>CSS</p>	 <p>Pandas</p>  <p>NumPy</p>  <p>Missingno</p>
 <p>협업 도구</p>	 <p>SQL</p>  <p>Json</p>  <p>Python</p>	 <p>matplotlib</p>  <p>Seaborn</p>  <p>NLTK</p>
 <p>Git</p>  <p>SVN</p>  <p>Google</p>	 <p>PyCharm</p>  <p>SQL Developer</p>  <p>Anaconda</p>	 <p>ScikitLearn</p>  <p>Scipy</p>  <p>statsmodels StatsModels</p>
	 <p>Flask</p>  <p>jQuery</p>  <p>Bootstrap</p>	 <p>Surprise</p>  <p>Xml</p>  <p>Ajax</p>
	 <p>SVD</p>  <p>ANCharts</p>  <p>ArmCharts</p>	





데이터

인공지능 프로젝트과정 이해와 산출물의 요소

- 데이터를 수집하기 위한 전략 수립
- 어떤 사이트에서 어떤 데이터를 수집할 것인가?
- 어떤 방법으로 어떤 도구를 이용해서 수집할 것인가?

공공기관의 빅데이터를 활용할 필요가 있음

MAIN

질병관리본부
KCDC

SUB

h-well
국민건강보험

식품안전나라

보건복지부

- 자연어처리
 - CSV 파일 형식, 문장 형식, 어노테이션이 있는 형식
 - 크기(MB), 행의 수 -> 임베딩한 후의 배열의 크기
- 이미지
 - 컬러영상, 이미지 개수, 크기(MB)
 - 어노테이션 형식(xml, json, txt) - 예시 캡처 포함
- 데이터
 - train, test
 - train, test, valid

“

질환 예측을 위한 분류분석용 데이터로 활용
(개인정보 및 질환진단 데이터)

”

“

식품 섭취에 따른 영양소 섭취량
산출 데이터로 활용

”

SK T-map 이용 데이터를 기반으로 구글맵 개인 데이터 크롤링

데이터 수집

데이터 내용

웹 크롤링을 통한 구글 맵 리뷰 데이터

SK 빅데이터 허브

2020년 4월 T map 사용 정보를 기반으로
시, 군, 구별 목적지 정보



- 결측치, 이상치 탐색 및 제거
- 표준화 또는 정규화
- 레이블 인코딩, 원-핫 인코딩
- 전처리 대상 열의 선택 기준 정의
- 다른 값을 대체할 경우 이유와 타당성 명시

01 원문 데이터

02 문서 정제

- ▶ 외부 링크 삭제
- ▶ 특수 기호 삭제
- ▶ 비속어 제거

03 자연어 처리

- ▶ 문장 분리 후
- ▶ 형태소 분석

04 키워드 및 연관어 추출

- ▶ 연관어 및 감성어 추출
- ▶ 분류 매핑

데이터 요약 통계량 탐색 후 데이터 전처리를 어떻게 할 것인가?



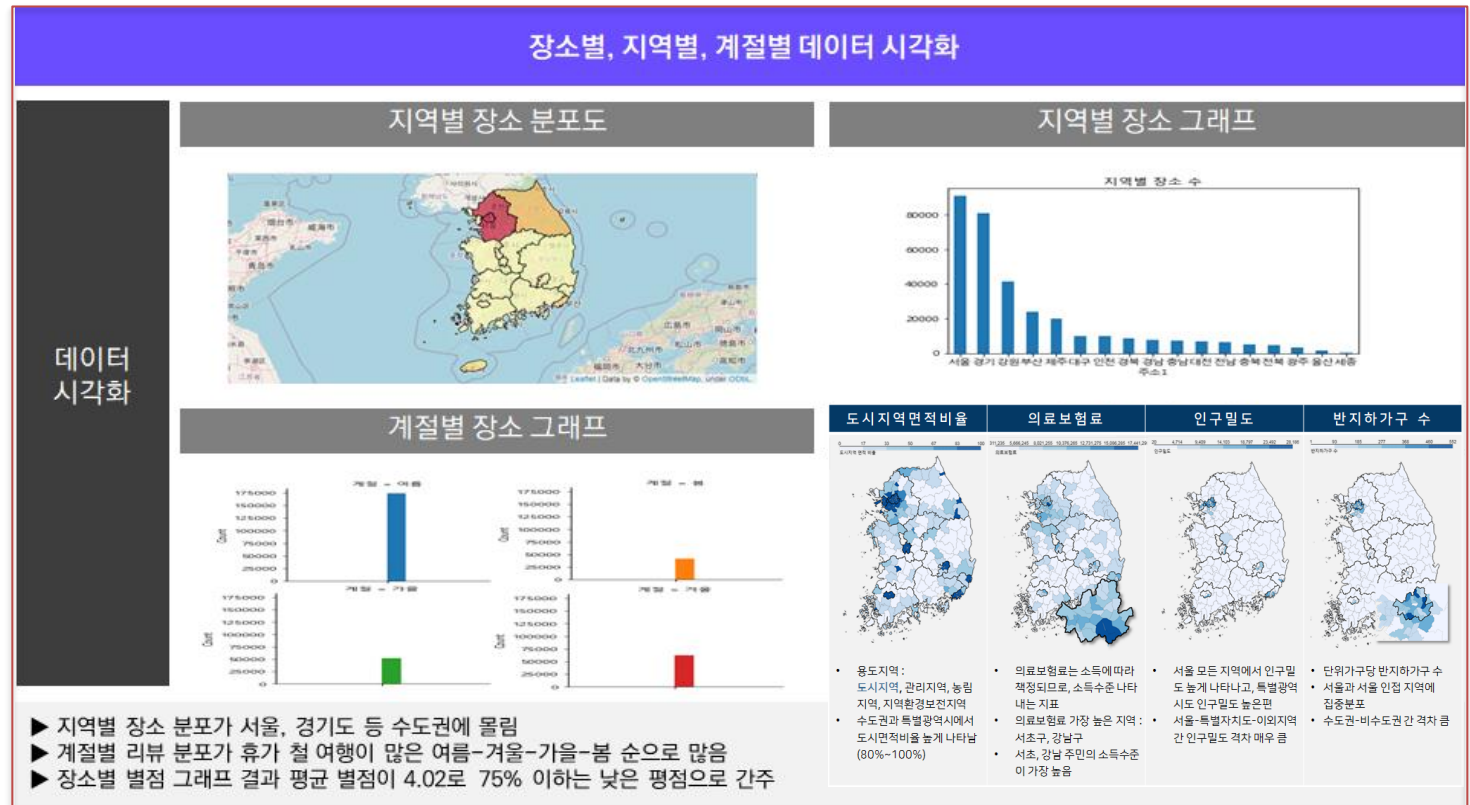
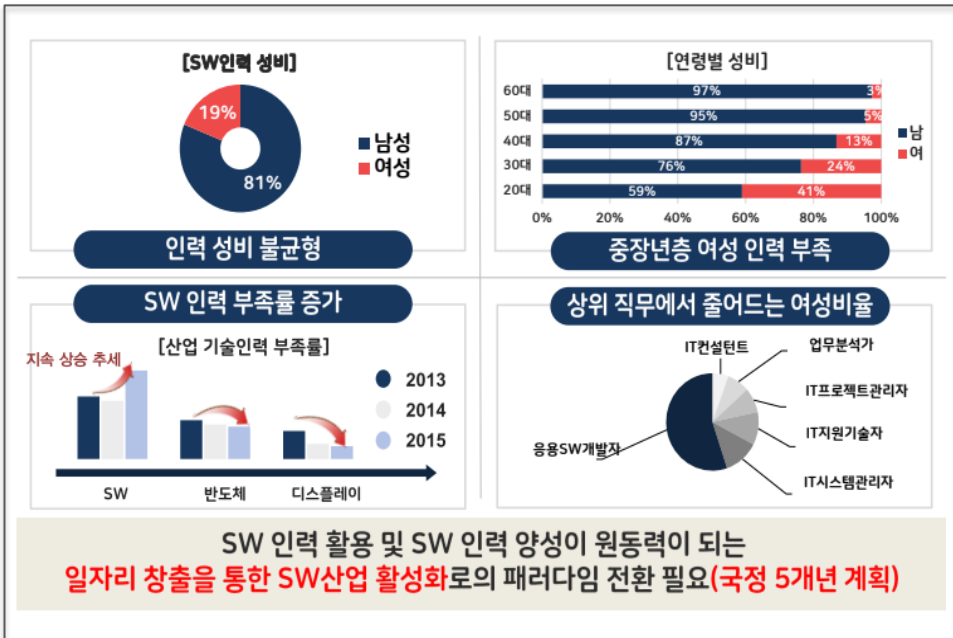


데이터 탐색

인공지능 프로젝트과정 이해와 산출물의 요소

- 탐색적 데이터 분석 수행
- EDA 결과를 시각화 할 필요가 있음

시각화 결과를 어떻게 정리할 것인가를 미리 생각해야 함.





프로젝트 수행 결과

인공지능 프로젝트과정 이해와 산출물의 요소



[프로젝트 수행 결과]는 프로젝트 결과물이 도출된 과정을 세부적으로 기록

- 예시는 하나의 사례로 간단하게 제시한 것이므로 프로젝트의 성격에 따라 보다 자세하게 기록하며, 결과를 서술하는 과정에서는 활용된 기술(구현 방법), 핵심기능, 검증 결과* 등을 상세히 기재한다.

* 예. 인공지능 객체탐지일 경우 IoU50 mAP 0.9, 빅데이터 직종의 경우 정확도, f1-score 등

- 프로젝트의 결과는 그 과정이 잘 드러날 수 있도록 가공 과정부터 활용까지 전체적인 프로세스를 확인할 수 있도록 단계별로 작성

* 첨부 자료 예시: 결과물 사진, 시연 동영상 등 프로젝트의 우수성이 드러날 수 있는 자료



프로젝트 수행 결과 - 머신러닝 기법 또는 딥러닝 모델

인공지능 프로젝트과정 이해와 산출물의 요소

- 머신러닝/딥러닝 기법 및 기본 이론에 대한 설명 추가
- 모델을 평가한 결과 기록
- 딥러닝의 경우 딥러닝 모델을 도식화 함

유사 모델을 비교 분석하는 것도 필요함

원-핫 인코딩 단어를 학습시키기 위해 벡터로 변환하는 작업 중에서 가장 간단한 방법으로, 0 또는 1로 변환합니다.

BUT 단어간의 유사도를 파악할 수 없습니다. 모든 벡터가 수직이므로 cosine similarity = 0
3개의 단어는 3차원의 벡터가 되므로 차원이 너무 커집니다.

임베딩 원-핫 인코딩의 단점을 보완한 것입니다. 원-핫 인코딩보다 저차원의 벡터를 생성합니다.

백터끼리의 유사도 구할 수 있습니다.
* 언어를 기계학적 공간 매핑하여 단어 사이에 있는 관계를

모델을 직접 구현하지 않았더라도 모델의 구조를 표현하고 설명할 수 있어야 합니다.

Embedding
[1,2]
[1,3]
[5,1]
[5,3]

VS

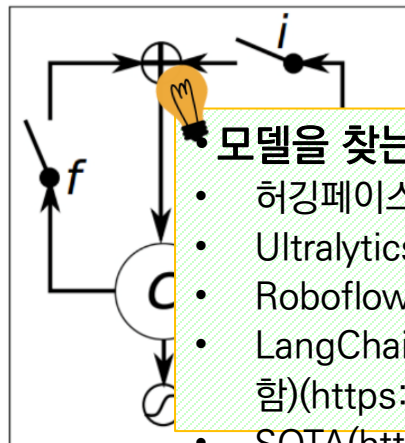
주변에 있는 단어로 중간에 있는 단어를 예측합니다

중간에 있는 단어를 가지고 중간에 있는 단어를 예측합니다.

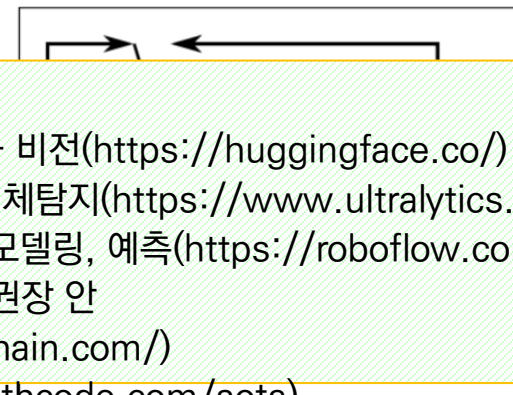


모델을 선택한 이유

- 모델의 구조를 표현해야 합니다.
- 구현한 소스코드 또는 summary() 결과를 활용하세요.
- 평가 방법에 대한 설명이 추가되어야 합니다.



(a) Long Short-Term Memory



(b) Gated Recurrent Unit

모델을 찾는 곳

- 허깅페이스: 자연어 모델 + 비전(<https://huggingface.co/>)
- Ultralytics-YOLOv8 : 객체탐지(<https://www.ultralytics.com/>)
- Roboflow : 어노테이션, 모델링, 예측(<https://roboflow.com/>)
- LangChain: API를 사용(권장 안 함)(<https://www.langchain.com/>)
- SOTA(<https://paperswithcode.com/sota>)

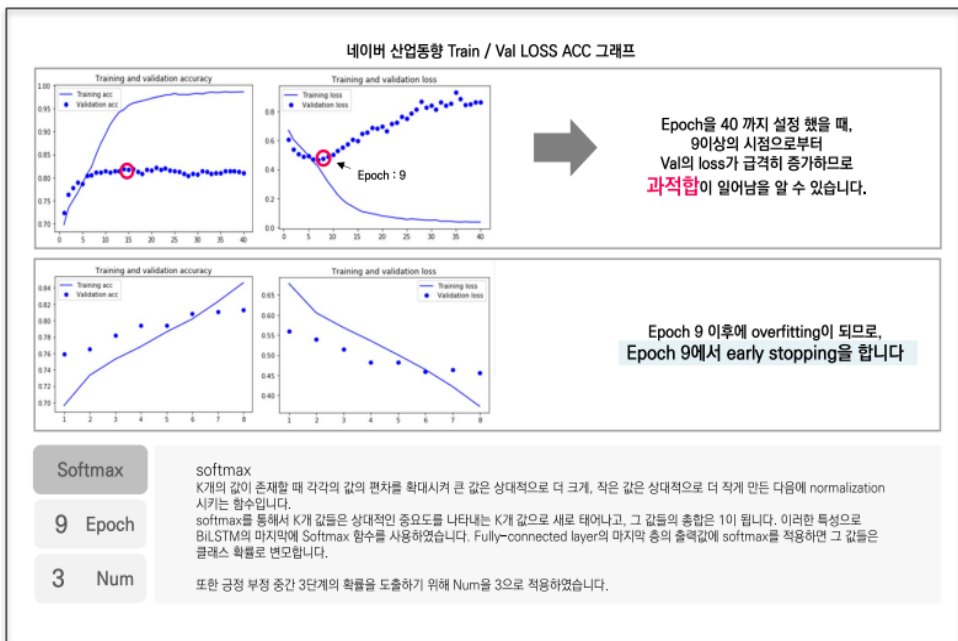


프로젝트 수행 결과 - 파인 튜닝 또는 모형 최적화

인공지능 프로젝트과정 이해와 산출물의 요소

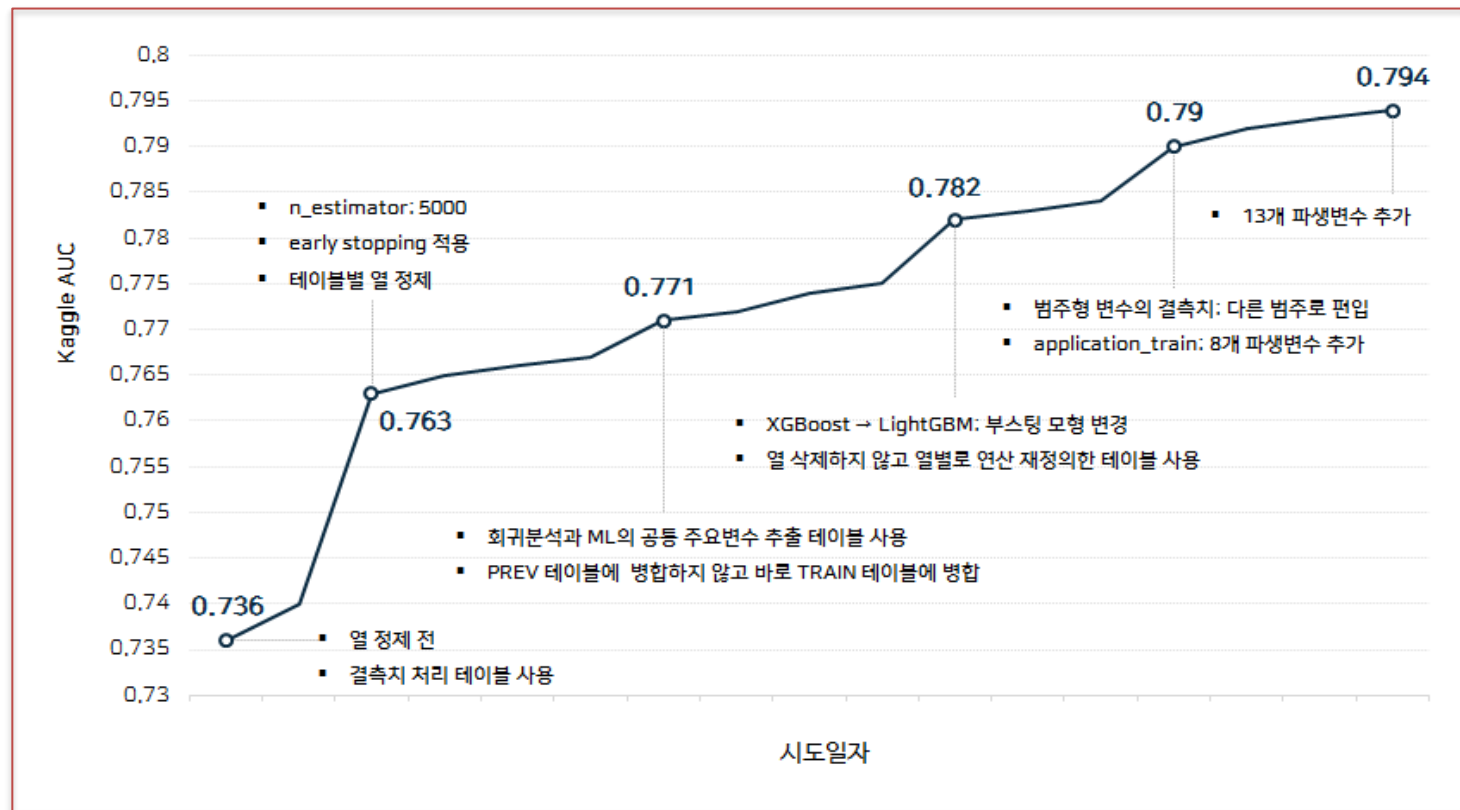
- 모형 최적화를 위해 수행한 내용 기록
- 파라미터 탐색
- 앙상블 모형 적용
- 과적합 여부 검사

최적화 전/후의 모형을 비교



모형의 평가 방법 및 평가 점수 개선을 위한 노력이 포함되어야 합니다.

- 모형의 수정 또는 하이퍼파라미터 조정 등 작업이 필요합니다.
- 모형을 비교할 수 있는 엑셀파일을 만들고 기록하세요.





프로젝트 수행 결과 - 모형 평가 비교

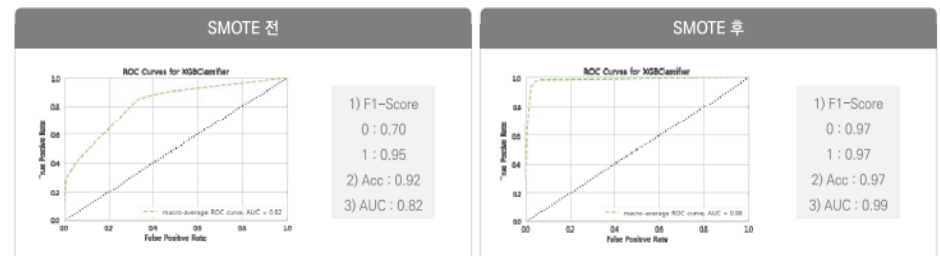
인공지능 프로젝트과정 이해와 산출물의 요소

- 모형을 평가하는 방법을 표시
- 평가방법을 선택한 근거와 이유 제시
- 교차검증 실시

교차검증을 통한 모형 평가가 필요할 수 있음



여러 모델에 교차검증을 실시하고, 얻어진 예측 결과값들을 합쳐 학습





프로젝트 수행 결과 - 모형별 평가표

인공지능 프로젝트과정 이해와 산출물의 요소

- 데이터의 크기와 알고리즘, 평가 점수를 기록해야 함
- 전처리 방법에 따른 평가점수
- 파생변수들 추가에 따른 평가점수
- 분석/예측 모형에 따른 평가점수

비교표를 제시해야 함

SMOTE 전

Decision Tree

	F1-Score	Acc	AUC
0	0.59	0.85	0.78
1	0.91		

Random Forest

	F1-Score	Acc	AUC
0	0.54	0.85	0.91
1	0.91		

MLPClassifier

	F1-Score	Acc	AUC
0	0.68	0.90	0.50
1	0.94		

CV Stacking Ensemble(XGBoost)

	F1-Score	Acc	AUC
0	0.70	0.92	0.82
1	0.95		

SMOTE 후

Decision Tree

	F1-Score	Acc	AUC
0	0.90	0.90	0.90
1	0.90		

Random Forest

	F1-Score	Acc	AUC
0	0.96	0.96	0.99
1	0.96		

MLPClassifier

	F1-Score	Acc	AUC
0	0.92	0.92	0.50
1	0.92		

CV Stacking Ensemble(XGBoost)

	F1-Score	Acc	AUC
0	0.97	0.97	0.98
1	0.97		

결론

SMOTE 전

F1-Score, Acc
RF < Stacking Ensemble
AUC
RF > Stacking Ensemble
→ Stacking Ensemble 채택

SMOTE 후

F1-Score, Acc
RF < Stacking Ensemble
AUC
RF > Stacking Ensemble
→ Stacking Ensemble 채택

최종 채택 Model

서비스 목적 : 성공(1) 예측

1에 대한 F1-Score를
SMOTE 전, 후로 비교했을 때 큰 차이가 없음

Data의 변형 위험을 감수하며
SMOTE를 사용하지 않고

SMOTE 전 Model 채택

모델링	Baseline	Library	Validation set	Without New_whale	Real-time Augmentation (Train)	Real-time Augmentation (Test)	Bounding -Box	Image padding + margin b-box	흑백 (grey)	Ensemble	Image Size (H/W)	batch_size	Score
#0													0.277
#1	CNN	Keras	x	X	X	X	X	X	X	X	100x100	64	0.290
#2	ResNet 18	Fast.ai	5개 미만증식 (복사)	X	O	O	X	X	X	X	384x384	64	0.565
#3	ResNet 18	Fast.ai		O	O	O	X	X	X	X	384x384	64	0.562
#4	ResNet 34	Fast.ai		X	O	O	X	X	X	X	384x384	64	0.588
#5	ResNet 34	Fast.ai		O	O	O	X	X	X	X	384x384	64	0.562
#6	ResNet 50	Fast.ai		X	O	O	X	X	X	X	382x382	*32	0.582
#7	ResNet 50	Fast.ai	사전증식 (랜덤증식)	O	O	O	X	X	X	X	382x382	*32	0.577
#8	ResNet 34	Fast.ai		X	O	O	O	X	X	X	384x384	64	Best score 0.614
#9	ResNet 34	Fast.ai		O	O	O	O	X	X	X	384x384	64	0.590
#10	ResNet 34	Fast.ai		O	O	O	O	X	X	X	384x384	64	0.588
#11	ResNet 34	Fast.ai		O	O	O	X	O	X	X	384x384	64	0.562
#12	ResNet 34	Fast.ai	Model Ensemble (3단계)	O	O	O	X	O	O	X	384x384	64	0.573
#???	ResNet 34	Fast.ai		X	O	O	O	X	O	X	384x384	64	
#???	ResNet 34	Fast.ai											



프로젝트 수행 결과 - 시연

인공지능 프로젝트과정 이해와 산출물의 요소

- 데이터의 크기와 알고리즘, 평가 점수를 기록해야 함
- 전처리 방법에 따른 평가점수
- 파생변수들 추가에 따른 평가점수
- 분석/예측 모형에 따른 평가점수

※ 별도파일로 첨부 가능



- 제출할 발표 영상은 5분 분량으로 만들어주세요.
- 시연 영상을 만들 때 자막을 포함하면 영상을 보는 사람이 더 이해하기 쉽습니다.



자체 평가 의견 또는 프로젝트 후기

인공지능 프로젝트과정 이해와 산출물의 요소

- ▶ [자체 평가 의견]은 프로젝트 결과물에 대한 프로젝트 기획 의도와의 부합 정도 및 실무 활용 가능 정도, 달성도, 완성도 등 자체적인 평가 의견과 느낀 점을 작성한다.
 - 개인 또는 우리 팀이 잘한 부분과 아쉬운 점을 작성한다.
 - * 예. 모델 평가 결과, 정확도가 00.00%로 정확도 향상을 위해 모델 추후 개선 필요
 - 프로젝트를 수행하면서 느낀 점이나 경험한 성과에 대하여 기재할 수 있으며, 경력 계획 등과 연관시켜 팀별 공통 의견 또는 개인 의견을 자유롭게 작성한다.