

## 동아리 활동보고서

활동명	2024 경기글로벌대전환포럼 AI 클래스룸 기획		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 15. ~ 2024. 10. 24.	비고	경기도 주최 국제 AI 포럼
활동목적	1. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다. 2. 세계적인 AI 석학과의 교류 기회 창출 ○ 연사자 - 피터 노빅(Peter Norvig, 前 구글 AI 연구 디렉터, 現 스탠포드 인간 중심 AI 연구소 특별 교육 연구원) - 스튜어트 러셀(Stuart Russell, 現 UC 버클리 전기공학 및 컴퓨터과학과 교수) 3. AI 클래스룸 기획 TF팀 구성을 통한 학생 주도의 대형 행사 기획 및 실행 역량 강화 4. 국제 AI 포럼 기획을 통한 경희대학교와 중앙동아리 KHUDA의 위상 제고		
활동내용 및 소감	<p>장소: 킨텍스 제2전시장 6홀 6A            주최·주관: 경기도, 킨텍스, 경기연구원            주제: 과거, 현재, 미래를 아우르는 인공지능 혁신 이야기            주요 연사: 피터 노빅, 스튜어트 러셀            참여 대상: AI에 관심 있는 누구나            행사 홈페이지: <a href="https://www.3gtf.com/">https://www.3gtf.com/</a>            특이사항: AI 클래스룸 기획 단계부터 실행까지 전 과정에서 KHUDA 가 주도적 역할 수행했습니다.</p> <p>활동 내용:            경기도에서 주최한 '2024 경기글로벌대전환포럼'에서 KHUDA가 AI 클래스룸 기획을 맡았습니다. "인공지능과 휴머니믹스"라는 큰 주제로, 글로벌 무대에서 AI가 어떻게 세상을 변화시키고 있는지에 대해 깊이 있는 논의가 펼쳐졌습니다. 인공지능 분야의 살아있는 전설, 세계적인 AI 석학인 스튜어트 러셀 교수(UC 버클리)와 피터 노빅 박사(Google Research Director)가 직접 참여하는 특별 토크쇼를 진행했습니다. AI 클래스룸에서는 인공지능의 미래와 인간 중심 AI의 방향성을 논의하는 자리로 마련했습니다.</p> <p>1) 학생 주도 AI 클래스룸 기획과 운영, 그리고 성공적인 마무리            경기도에서 주최한 '2024 경기글로벌대전환포럼'에서 KHUDA가 AI 클래스룸을 기획했습니다. 학생들이 주도적으로 대형 행사를 기획하고 실행할 수 있는 역량을 강화하기 위해 AI 클래스룸(AI Classroom) 기획 TF팀을 구성하였습니다. TF팀은 행사의 기획부터 운영까지 전 과정을 주도하였으며, 행사의 모든 세부사항을 직접 기획하고 실행함으로써 기획력과 실행력을 키울 수 있는 기회가 되었습니다.</p>		

AI 클래스룸은 미국 대학의 수업 방식을 재현하는 형식으로 진행했으며, 참가 학생들은 영어로 진행되는 강의와 토론을 경험하면서 글로벌 AI 트렌드를 학습하는 기회를 가졌습니다. 이를 통해 참가자들은 인공지능에 대한 깊은 이해를 높이는 동시에 국제적인 감각과 영어 토론 능력을 함양할 수 있었습니다.

## 2) 국제 AI 포럼 기획 및 KHUDA의 역할 확대

이번 AI 클래스룸 행사는 단순한 강연을 넘어 경희대학교와 중앙동아리 KHUDA의 AI 분야 위상을 높이는 중요한 계기가 되었습니다. 본 행사를 통해 우리 동아리는 학생들이 스스로 학술적 주제를 선정하고, 연구 및 발표 자료를 제작하며, 영어로 진행되는 토론을 주도할 수 있는 역량을 갖추었음을 증명해냈습니다.

행사는 KHUDA 회원들의 적극적인 참여로 성공적으로 진행되었습니다. 김재욱 학생이 모더레이터 역할을 맡아 행사의 원활한 진행을 이끌었으며, 홍민혁 학생은 '인공지능의 여정: 과거, 현재, 그리고 미래'를 주제로 첫 번째 발표를, 임소영 학생은 '인간 중심의 AI: 기술과 윤리의 조화'를 주제로 두 번째 발표를 진행하였습니다. 또한 KHUDA 회원들이 직접 영어로 모든 진행과 발제를 맡으며 행사의 수준을 높이는 데 기여하였습니다.

## 3) 행사 개요 및 진행 방식

AI 클래스룸은 인공지능의 과거, 현재, 미래를 중심으로 구성된 심도 있는 논의의 장이었습니다. 사전 기획 단계에서 TF팀은 최신 AI 트렌드와 주요 연구 주제를 분석하여 발표 주제를 선정하였고, 원활한 토론 진행을 위해 다양한 질문을 준비하였습니다. 또한 패널 Q&A 세션을 통해 참가자들이 세계적인 AI 연구 동향을 직접 접할 수 있도록 하였습니다.

행사는 전 과정이 영어로 진행되었으며, 참가자들은 사전 준비한 질문을 바탕으로 글로벌 수준의 학술적 토론을 경험할 수 있었습니다. 이번 행사를 통해 학생들은 AI 기술 및 윤리 문제에 대한 통찰력을 기르는 동시에, 영어 발표 및 토론 능력을 향상시킬 수 있었습니다.

## 타임라인 및 주요 토론 내용

타임라인	진행 내용	진행자/발표자
14:00 ~ 14:05	개회	정유진
14:05 ~ 14:10	주제발표1: '인공지능의 과거, 현재, 그리고 미래'	홍민혁
14:10 ~ 14:35	<b>토크쇼 - 패널 Q&amp;A</b> Q1. 현재의 생성형 AI 트렌드가 앞으로의 AI 연구 방향에 어떤 영향을 미칠 것으로 예상하시나요? Q2. AI가 인간 수준의 지능을 갖추기 위해 극복해야 할 가장 큰 기술적 장벽은 무엇이라고 생각하십니까?	김재욱
14:35 ~ 14:40	주제발표2: '인간 중심의 AI: 기술과 윤리의 조화'	임소영
14:40 ~ 15:00	<b>토크쇼 - 패널 Q&amp;A</b> Q3. AI의 발전이 현재의 기업 문화와 조직 구조에 어떤 변화를 가져올 것으로 예상하시나요?	김재욱

	Q4. '인간 중심의 AI'를 실현하기 위해 AI 시스템 설계 시 고려해야 할 핵심 원칙들은 무엇일까요?	
15:05 ~ 15:10	폐회	정유진

AI 클래스룸을 성공적으로 마치면서, 앞으로도 AI 관련 포럼이 열린다면 우리 동아리가 대학생 주도 행사를 맡아 기획하고 운영할 수 있도록 꾸준히 노력해 나가려 합니다. 해외 AI 전문가 및 연구진과의 교류 기회를 넓히고, 동아리원들이 글로벌 네트워크를 쌓을 수 있는 방향을 계속해서 고민해나가겠습니다.

또한 정규 세션 외에도 AI 윤리와 산업 응용에 대한 논의를 더 현실적인 시각에서 다룰 필요가 있다는 점을 깨달았습니다. 그래서 앞으로 특강 세션을 추가로 운영할 기회가 생긴다면, AI 기술이 실생활에서 어떻게 활용될 수 있는지에 대한 이해를 넓히는 데 초점을 맞추고자 합니다. 이런 활동을 통해 우리 동아리는 학생들이 자율적으로 AI를 연구하고 배우는 문화를 조성하고, 학문적 탐구와 실무 역량을 동시에 키울 수 있도록 돕고 싶습니다.

KHUDA가 단순한 교내 AI 학술 동아리를 넘어, 대한민국을 대표하는 AI 커뮤니티로 자리 잡을 수 있도록 앞으로도 한 걸음씩 나아가겠습니다.

#### 소감:

**강\*\*:** 이번 AI 클래스룸에 참여하면서 단순한 이론 학습을 넘어, 실전에서 AI를 어떻게 다뤄야 하는지 고민할 수 있는 뜻깊은 시간이었습니다. 더 나은 세션을 기획하고자 많은 고민을 했던 과정에서, 수많은 동아리원들이 적극적으로 질문하고 토론에 참여하는 모습을 보면서 보람을 느꼈습니다. 단순히 공부하는 것에서 끝나는 것이 아니라, 지식을 공유하고 함께 성장하는 것이 얼마나 중요한지 깨달았던 시간이었습니다. 다음에도 이런 기회가 있다면 더 나은 세션을 만들기 위해 도전해보고 싶습니다.

**김\*\*:** TF팀으로 참여했지만, 오전에 교과목 시험이 있어 행사 당일에 직접 참여하지는 못했습니다. 대신 라이브로 유튜브에 AI 클래스룸이 중계되어, 같은 공간에 함께 있는 기분이 들었습니다. 함께 기획한 내용을 세계적인 AI 석학 두 분(책에서만 보던 분들.. 직접 보고 이야기 나누신 분들 정말 부럽습니다ㅜㅜ) 앞에서 두려움 없이 잘 마무리해내는 우리 동아리원들이 정말 멋지고 자랑스럽게 느껴졌습니다!!

**이\*\*:** 클래스룸을 통해 다양한 참가자들과 함께 AI에 대해 논의하면서, 기술적인 지식뿐만 아니라 서로 다른 시각과 접근 방식을 배울 수 있었습니다! 제 개인적인 희망 사항이지만 해외 연구진과의 협업이 가능하다면 어떨까 하는 생각도 들었고, 앞으로 우리 동아리가 글로벌 AI 네트워크의 중심이 될 가능성이 충분하다는 확신이 들었습니다!! 단순한 강의나 세미나를 넘어서서, 실제로 의미 있는 연구와 프로젝트로 발전시킬 수 있도록 계속해서 노력해야겠다고 다짐했습니다. 모두들 고생 많으셨습니다ㅎㅎ

**정\*\*:** 개인적으로 AI 기술의 발전만큼이나 윤리적인 문제도 중요하다고 생각해 왔는데, 이번 클래스룸을 통해 그 고민이 더욱 깊어졌습니다. 저는 그 중에서도 AI가 산업에 적용될 때 발생할 수 있는 문제와 해결 방안에 대한 논의가 인상적이었습니다. 기존에는 기술적인 부분에만 관심

이 많았는데, 앞으로는 AI가 사회적으로 미칠 영향까지 고려하며 공부해야겠다는 생각이 들었습니다. 이런 기회를 만들어 준 우리 KHUDA 동아리와 TF팀원들에게 깊은 감사를 드립니다.

최\*\*: 처음에는 AI 클래스룸을 단순히 지식을 배우는 자리라고 생각했습니다. 그런데 막상 진행해보니 '배움을 나누는 즐거움'이 얼마나 큰지 깨달을 수 있었습니다. 질문을 준비하면서 많은 자료를 탐색하고, 또 자연스럽게 더 깊이 공부하게 되었습니다. 행사 자리에서 일반 참여자들 과도 자유롭게 의견을 나누면서 이전에는 생각조차 하지 못했던 새로운 시각을 배울 수 있었습니다. 혼자서 공부할 때와는 다르게 함께 고민하고 토론하는 과정 자체만으로도 정말 의미 있었습니다. 앞으로도 이런 활동이 많아졌으면 좋겠고, 동아리 활동을 더욱 적극적으로 참여해야겠다고 다짐했습니다!!

활동사진  
(3매 이상)













동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

## 동아리 활동보고서

활 동 명	제1회 KHUDA 데이터톤		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 11. 09. 9:00 ~ 21:00	비고	
활동목적	1. KHUDA 회칙 제1장 제3조 [목적] 학습한 내용을 바탕으로 실용성 있는 프로젝트를 진행하여 실무 및 연구 역량을 증진한다. 2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.		
활동내용 및 소감	<p>장소 : 전자정보대학 211-1호</p> <p>주최 : KHUDA</p> <p>후원 : 현대모비스</p> <p>대상 : 데이터를 활용해 문제 해결 능력을 기르고, 협업을 통해 데이터 분석 역량을 높이고 싶은 대학생 누구나</p> <p>연사 : 이동훈 (코로나맵 개발자)</p> <p>참여 인원 : 총 64명 (참가자 53명, 운영진 11명)</p> <p>참여 학교: 경희대학교, 서울대학교, 한국외국어대학교, 숭실대학교, 인하대학교 등 7개교</p> <p><b>활동내용 :</b></p> <p><b>1) 기획 의도</b></p> <p>개발 중심의 해커톤은 많지만, 데이터 분석에 중점을 둔 데이터톤은 드물기에, 학생들에게 실질적인 데이터 분석 경험을 제공하고 타 학교 학생들과의 교류를 통해 학문적 협업과 실무 역량 강화의 기회를 마련하고자 데이터톤을 개최하게 되었습니다. 본 행사의 목표는 데이터 활용 문제 해결 능력 향상과 협업을 통한 데이터 분석 역량을 강화하는 것이었습니다. 또한, 현대모비스의 후원과 코로나맵 개발자 이동훈님의 강연을 통해 참가자들에게 보다 깊이 있는 경험을 제공하였습니다.</p> <p><b>2) 행사 개요</b></p> <p>데이터톤은 다양한 학교의 데이터 분석에 관심 있는 학생들이 팀을 이루어 참여하는 방식으로 진행되었습니다. 참가자들은 주어진 데이터셋과 추가 데이터를 활용하여 데이터 분석 인사이트를 도출하고 활용 방안을 제시하는 것을 목표로 하였습니다. 결과물은 웹이나 앱 형태가 아닌, 데이터 분석 결과를 기반으로 평가되었습니다.</p> <p>총 7개 대학의 학생이 참여하였으며, 5~7인 1팀으로 총 9팀으로 구성했습니다.</p> <p><b>1. 주요 혜택</b></p> <ul style="list-style-type: none"> <li>- 참가자 전원: 점심 및 저녁 식사, 간단한 다과 제공, 다양한 데이터 분석 분야 학생들과의 교류 및 네트워킹 기회 제공</li> </ul>		

- 수상자: 최우수상(30만원), 우수상(10만원, 2팀 선정)

## 2. 참가비 및 운영비

- 참가비: 10,000원
- 운영비: 참가비(10,000원) + 지원금(15,000원) = 25,000원/인

## 3) 데이터톤 진행 방식

데이터톤의 핵심 키워드는 “비빔”이었으며, 5가지의 세부 주제는 행사 당일 공개되었습니다. 참가자들은 팀별로 2개 이상의 키워드(비빔 재료)를 활용하여 데이터 분석을 수행했습니다. 최종적으로 데이터 분석을 통한 인사이트 도출 및 활용 방안을 제시하는 방식으로 진행되었습니다.

### 1. 5대 키워드 소개

키워드	선정 이유
금융/경제 (Finance /Economy)	디지털화와 함께 금융 산업의 변화 속도는 가파르게 증가하고 있습니다. 핀테크, 블록체인, 디지털 자산 등의 기술 발전은 현재 금융 생태계 전반에 걸쳐 영향을 미치고 있으며, 이에 따라 최근 데이터 분석과 AI를 활용한 금융 솔루션이 널리 활용되고 있어 첫 번째 키워드로 선정합니다.
문화/여가 (Culture/ Leisure)	엔데믹 이후, 사람들은 팬데믹 동안 제한되었던 문화/여가 생활에 대한 수요가 급증하고 있습니다. 대규모 행사 및 해외 여행이 재개되면서 문화예술, 레저 활동, 국내외 여행에 대한 관심이 폭발적으로 증가하는 흐름을 반영해 두 번째 키워드로 선정합니다.
기후 (Climate)	2024년에 들어서 극단적인 폭염, 스콜성 폭우 등 이상 기후 현상이 빈번해지며, 기후 위기와 분석의 중요성이 더 이상 미래의 일이 아니라는 것을 느낄 수 있었습니다. 기후 예측 모델을 강화하거나, 기후 리스크를 모니터링하고, 재해 대비 계획을 수립하는 추세에 따라 세 번째 키워드로 선정합니다.
한국어 (Korean Language)	K-POP, K-콘텐츠 등 높아진 한국 문화의 관심에 따라 한국어에 대한 글로벌 관심도 덩달아 크게 증가하고 있습니다. 또한, 한국어가 언어적으로 갖는 독창적이고 유일무이한 특징이 “데이터 분석” 분야에서 갖는 잠재력이 높다고 판단하여, 데이터톤 네 번째 키워드로 선정합니다.
식품/농축산 (Food /Agriculture)	최근에는 스마트팜, 로봇 농업, 수직 농업 등 기술을 활용한 혁신적인 농업 방식이 주목받고 있습니다. 지속 가능한 식량 생산과 농업 혁신은 기후 변화와 환경 보호, 우리의 미래와 직결된 중요한 부분이며 그에 따른 관심도 갈수록 높아지는 추세입니다. 해당 흐름을 이어 마지막 키워드로 선정합니다.

### - 예시 분석 주제:

- 유사한 이름의 기업 주가 상관성 분석 (금융/경제 + 한국어)
- 날씨에 따른 식당 웨이팅 시간 예측 (기후 + 식품)



2. 데이터 제공 방식

아래 이미지와 같이 데이터셋은 운영진이 각 키워드에 맞추어 미리 준비한 데이터셋을 제공하였습니다. 제공된 데이터 외에도 타 기관 데이터를 자유롭게 활용할 수 있도록 하였습니다. 또한 참가자들이 데이터 수집에 용이하도록 키워드별 데이터셋 참고 사이트 리스트를 제공하였습니다.

2. 키워드별 제공 데이터셋 리스트

해당 데이터셋은 데이터톤 5개의 키워드에 맞춰 운영진 측에서 준비한 데이터입니다.  
데이터 분석 시 해당 데이터셋을 사용하셔도/사용하지 않으셔도 되며, 이와 추가적인 데이터는 아래 (3)에서 첨부한 참고 사이트를 확인해 구하시면 됩니다.

표

키워드	Aa 데이터셋 이름	압축 파일	데이터셋 소개
금융/경제	집계구 단위 업종별 지역화폐 거래 데이터 (부산시)	집계구 단위 업종별 ...	부산시 지역화폐(동백전) 행정동 단위, 업종별 거래 데이터입니다. - (1) 집계 차원 : 기준일자(YYYYMMDD), 기준시간, 시도명, 시군구명, 행정동명, 집계구코드, 집계구중심 경도, 집계구중심 위도, 카드사 명칭, 업종코드(BC카드), 업종명칭(BC카드), 생년(YYYY), 성별(M or F) (2) 집계 항목 : 거래횟수, 총거래금액
금융/경제	행정동 단위 업종별 지역화폐 거래 데이터 (부산시)	행정동 단위 업종별 ...	부산시 지역화폐(동백전) 행정동 단위, 업종별 거래 데이터입니다. - (1) 집계 차원 : 기준일자(YYYYMMDD), 기준시간, 시도명, 시군구명, 행정동명, 행정동코드, 카드사 명칭, 업종코드(BC카드), 업종명칭(BC카드), 생년(YYYY), 성별(M or F) (2) 집계 항목 : 거래횟수, 총거래금액
금융/경제	전통시장별 지역화폐 거래량, 금액 데이터 (부산시)	전통시장별 지역화...	부산시 지역화폐(동백전) 전통시장별 단위, 업종별 거래 데이터입니다. - (1) 집계 차원 : 기준일자(YYYYMMDD), 기준시간, 전통시장 명칭, 전통시장 주소, 카드사 명칭, 업종코드(BC카드), 업종명칭(BC카드), 생년(YYYY), 성별(M or F) (2) 집계 항목 : 거래횟수, 총거래금액
문화/여가	관광 KVQA 데이터 (경기도)	관광 데이터.zip 219-1.관광 KVQA ...	지식정보(텍스트)로 이루어진 다중쌍데이터셋으로 관광지, 문화시설, 음식점, 레포츠, 숙박, 편의 오락에 해당하는 데이터를 가지고 있습니다. 원본파일에 해당하는 <219-1.관광 KVQA 데이터(경기)>의 경우 json형식으로 되어있고, 각각의 시설의 고유 ID 와 해당 시설의 정보들이 포함되어있습니다. <관광 데이터> 파일은 위의 json파일을 csv파일로 가공한 파일에 해당합니다.
문화/여가	문화예술 관람 현황 및 희망 분야 설문조사 데이터	문화예술 관람 현황 ...	온라인(PC모바일) 소비자 서베이 데이터(2021.11월부터 매주 수집) • 문화예술 관람 분야 8종(무용 공연, 문학 행사, 미술관전시관, 연극뮤지컬 등) 중 지난 동시기 대비 관람횟수가 증가한감소한 분야는 무엇이고, 향후 1~2년 이내 지출을 늘려서라도 하고 싶은 분야가 무엇인지에 대해 응답한 설문조사 데이터입니다. • 본 데이터 셋에는 응답자 특성(성별, 연령대, 거주지역, 소득수준 등)이 포함. • 데이터 출처 ㈜ 컨슈머인사이트 정기 기획조사 '여가문화체육 조사'
기후	기상 예보 데이터 (2022-23) [부산 중앙동]	부산중구중앙동_기...	단기예보는 예보기간과 구역을 사-공간적으로 세분화하여 발표하는 예보입니다. 지역별 시간별 차이로 인한 수 요자의 불편을 최소화하기 위해 전국을 5km*5km 간격의 격자로 나누어 읍, 면, 동 단위의 행정구역 중심으로 상세한 날씨를 제공합니다.

3. 키워드별 데이터셋 참고 사이트

해당 사이트들은 키워드에 맞춰 데이터를 구하는데, 데이터톤 참여자 여러분들의 편의를 위해 마련해둔 링크 리스트입니다.  
데이터 분석 시 해당 데이터 사이트 이외의 공간에서 데이터를 수집하셔도 무방하다는 점 공지드립니다.

표

키워드	Aa 사이트 이름	링크	참고사항
금융/경제	금융 빅데이터 플랫폼	bigdata-finance.kr/dat_ist.do	상세조건을 "무료"로 설정하면 데이터셋을 내려받을 수 있습니다.
문화/여가	문화 빅데이터 플랫폼	bigdata-culture.kr/big_ieYn=N	상세조건을 "무료"로 설정하면 데이터셋을 내려받을 수 있습니다.
기후	기상청 날씨데이터 서비스 : 기상 자료 열기	data.kma.go.kr/cmm_ain.do	날짜, 지역별 기상 관측 데이터 및 API를 무료로 이용하실 수 있습니다.
기후	환경 데이터 마켓	bigdata-environment.kr/use_ist.do	상세조건을 "무료"로 설정하면 데이터셋을 내려받을 수 있습니다.
한국어	한국어 데이터셋 깃허브 모음	github.com/son_v-file	한국어 분야와 관련된 데이터셋을 정리해둔 깃헙 레포가 있어 공유드립니다.
식품/농축산	농식품 빅데이터 거래소	kadx.co.kr/	한국농수산식품유통공사에서 제공해주는 데이터셋으로, 농식품 공공데이터를 무료로 다운받으실 수 있습니다.
식품/농축산	농림축산식품 공공데이터 포털	data.mafra.go.kr/main.do	농림축산식품부에서 제공해주는 데이터셋, 농림축산 + 식품/정책/유통소비 등 포함
공통	AI Hub	aihub.or.kr/aih_nu=100	-
공통	공공데이터포털	data.go.kr/tcs_ist.do	-
공통	서울시 열린데이터 광장	data.seoul.go.kr/dat_Flag=M	-
공통	경기도 공공데이터 개방포털	data.gg.go.kr/por_age.do	-
공통	Kaggle	kaggle.com/datasets	-

#### 4) 행사 진행 일정

타임라인	일정 상세
09:00 ~ 09:10	참가자 등록 및 입장
09:10 ~ 09:40	KHUDA 데이터톤 오리엔테이션 및 주제 공개
09:40 ~ 10:00	현대모비스 후원사 홍보 세션
10:00 ~ 13:00	오전 활동 및 중간 과제 제출
13:00 ~ 14:00	점심 식사 제공
14:00 ~ 15:00	연사 강연: 이동훈 (코로나맵 개발자)
15:00 ~ 19:00	오후 활동 및 최종 과제 제출
19:00 ~ 19:30	저녁 식사 제공 및 우수팀 선정 투표 진행
19:30 ~ 20:00	우수팀 발표
20:00 ~ 21:00	심사 결과 발표 및 시상식, 사진 촬영

#### 5) 연사 강연

이번 데이터톤에서는 유명 방송 '유클즈'에도 출연한 코로나맵 개발자 이동훈님을 초청하여 강연을 진행하였습니다. 강연에서는 데이터 분석 및 활용, 그리고 개발이 어떻게 이루어지는지에 대한 실제 사례를 공유하며, 참가자들이 데이터 분석 실무에 대한 이해를 높일 수 있도록 하였습니다.



## 6) 평가 방식

참가자들은 발표 자료를 확인한 후 개별 투표를 진행하였으며, 이를 통해 상위 3개 팀을 선택한 뒤 본격적인 평가가 이루어졌습니다. 선정된 팀들은 Google Meets를 통해 평가자들에게 직접 발표를 진행하였으며, 평가자들은 발표 내용을 듣고 심사를 진행하는 방식으로 평가가 이루어졌습니다.

- 심사 위원: 소프트웨어융합학과 성무진 교수님, KAIST 금융공학 박사과정 4년차 최인수, KHUDA 3~4기 회장 임소영 학생
- 평가 기준은 운영진 내부 회의를 거쳐, 데이터톤 개최의 취지에 맞게 아래와 같이 평가 항목을 균형 있게 반영할 수 있도록 설정하였습니다.

## 세부 평가 기준

평가 항목	내용
이해도	데이터에 대한 이해도가 얼마나 높은지
논리성	결과 도출까지의 논리성, data-driven 주장의 일관성
적합성	아이디어 구현 및 해석이 얼마나 적합한지
발전가능성	제시한 아이디어와 해결 방안이 실현 가능한지
창의성	아이디어나 문제 접근 방식이 얼마나 참신한지

## 7) 팀별 분석 주제 및 결과

팀명	분석 주제	선택 키워드
1팀	우울증과 관련된 다양한 원인 분석과 그 영향	금융/경제, 기후, 한국어
2팀	날씨에 따른 치킨 판매량 및 공급 영향 정도 확인	금융/경제, 기후, 식품/농축산
3팀	데이터 분석 기반 지역 특성과 방언을 반영한 맞춤형 축제 및 마케팅 시스템 개발	금융/경제, 문화/여가, 기후, 한국어, 식품/농축산
4팀	입점 조건에 따른 업장 종류, 업장명 및 농산물 추천	금융/경제, 한국어, 식품/농축산
5팀	과잉생산 농산물을 활용한 푸드업사이클링 제품 제안 및 네이밍	금융/경제, 문화/여가, 기후, 한국어, 식품/농축산
6팀 (우수상)	성공적인 지역 축제 준비를 위한 서비스	금융/경제, 문화/여가, 기후, 한국어, 식품/농축산
7팀 (우수상)	매출 데이터와 트렌드 분석을 통한 편의점 히트 상품 예측 및 신제품 추천	금융/경제, 문화/여가, 기후, 식품/농축산
8팀	KOSPI 지수에 영향을 미치는 요인 분석	금융/경제, 기후, 식품/농축산

9팀  
(최우수상)

ESG K - CULTURE ETF 상품 개발

금융/경제, 문화/여가,  
기후, 한국어, 식품/농축산

8) 운영 및 준비 과정

행사의 원활한 운영을 위해 데이터톤 기획 TF팀을 구성하였습니다. TF팀 구성을 통해 학생 주도의 대형 행사 기획 및 실행 역량을 강화할 수 있는 기회를 제공했습니다. 행사를 기획해보고 싶은 누구나 지원할 수 있었으며, 각 팀은 다음과 같이 역할을 분배하였습니다.

- **행사 기획팀:** 전체 행사 기획 및 일정 조정, 데이터셋 준비 및 분석 주제 설정
- **참가자 관리 및 대외 협력팀:** 참가자 모집 및 커뮤니케이션, 후원사 협력
- **운영 및 현장 관리팀:** 행사 당일 현장 운영 및 관리, 식사/음료 및 행사 필요 물품 준비

9) 결론 및 향후 계획

제1회 KHUDA 데이터톤은 참가자들에게 실질적인 데이터 분석 경험을 제공했을 뿐만 아니라, 타 학교 학생들과의 네트워킹을 촉진하는 의미 있는 행사였습니다. 참가자들은 제한된 시간 안에서 문제를 해결하며 협업의 중요성을 직접 체험할 수 있었고, 데이터 분석 실력을 한층 성장시킬 수 있었다는 긍정적인 평가가 많았습니다.

이번 데이터톤은 단순한 행사에 그치는 것이 아니라, KHUDA의 역사적인 첫걸음으로서 앞으로 더욱 발전시켜 나갈 중요한 초석이 될 것이라 생각합니다. 향후에는 더욱 다양한 주제와 심화된 분석 기회를 제공하여 KHUDA 회원뿐만 아니라 데이터에 관심 있는 주변 지역 학생들까지 함께 성장할 수 있는 장을 마련하고자 합니다. 이를 통해 데이터 분석 문화를 더욱 확산시키고, 지속적인 발전을 이루어 우리 학교의 명성을 높이는 데 기여하고 싶습니다.

소감 :

[참가자 소감]

**임\*\*:** 힘들었지만 많은 데이터를 만져보고 시각화하는 과정이 재밌었어요! 결과가 예상대로 잘 나오지 않아 아쉬운 부분도 있었지만, 정말 좋은 경험이었습니다. 시간이 조금 더 많았다면 제대로 분석할 수 있었을 것 같습니다. 다음에는 더 완성도 있는 분석을 진행하고 싶습니다!

**김\*\*:** 오전 9시부터 오후 9시까지 진행했는데, 막상 하다보니 시간이 너무너무너무 짧게 느껴졌습니다. 하지만 주어진 시간 내에 완성도 있는 결과물을 내기 위해 팀원 모두가 노력했던 보람찬 프로젝트였습니다.

**홍\*\*:** 개발에 중점을 둔 해커톤은 정말 많은데, 데이터톤은 많지 않은 실정입니다. 이렇게 빠른 시간 내에 데이터 분석 실력을 향상시킬 수 있는 좋은 기회를 주신 운영진분들께 감사합니다.

**윤\*\*:** 진행 중에 예상과 달리 흘러가 주제 변경에 어려움이 있었지만, 하루 종일 데이터 분석을

경험하며 부족함 점을 깨달을 수 있어 유익한 시간이었다.

신\*\*: 책으로 공부하던 내용들을 토론과정에 떠올리며 적용해보는 것이 흥미로운 경험이었고 추후 다른 프로젝트를 위해 평소에 공부를 좀 해야겠다.

김\*\*: 기회가 된다면 또 참여하고 싶네요. 좋은 사람들과 길고 힘들지만 재밌는 시간 보낼 수 있어서 좋았습니다. 주제 선정 시간이 짧은 감이 있어서, 다음번에는 무박 2일로 진행해도 좋을 것 같습니다.

서\*\*: 데이터 분석에 대해서 아직도 모르고 부족하다는 것을 다시 깨달을 수 있었습니다. 그래도 행사 재밌게 즐긴 것 같아서 좋았습니다. 나중에 기회가 된다면 KHUDA에서 활동하면서 많이 배우고 싶습니다.

임\*\*: 재밌습니다. 여러 주제를 엮는다는게 처음에는 다소 낯설었지만 오히려 여러 주제를 섞음으로써 예상치 못하는 인사이트가 나올 수 있다는 것을 배웠습니다.

조\*\*: 경희대학교 학생이 아니라 행사를 위해 수원까지 왔는데, 먼 걸음이 전혀 아깝지 않았습니  
다. 데이터 분석은 정말 어렵다고 느꼈고 다시 공부를 해야겠다라는 생각이 들었습니다. 데이터분석 프로젝트에 처음으로 참여하여 많은 긍정적인 자극을 받고 갑니다. 이런 기회를 만들어주셔서 감사합니다. 팀원들도 다들 너무 잘해주셔서 재밌게 했습니다.

#### [TF팀 운영진 소감]

정\*\*: 이번 학기까지 현대모비스 SW 지원 동아리로 선정되어 행사 운영에 있어 금전적인 부담이 크지 않았습니다. 그러나 앞으로 데이터톤을 더욱 발전시키기 위해서는 재정적인 부분을 보다 면밀히 고민할 필요가 있습니다. 특히 경희대학교 외 타 학교 학생들의 참여를 유도하려면 참가비를 높게 책정하기 어려운 반면, 일정 수준 이상의 상금이 필요하다는  
점이 중요한 과제입니다.

또한 이번 행사에서는 주제 선정 시간이 짧았다는 아쉬운 평가도 있었습니다. 향후에는 개발 해커톤처럼 무박 2일 일정으로 진행할 경우, 식사 및 간식비 등의 운영 비용이 더욱 증가할 것으로 예상됩니다. 다음 기수 운영진들이 지속 가능한 운영 방안을 마련하여 데이터톤을 더욱 발전시켜 나가면 좋겠습니다.

김\*\*: KHUDA 신규 회원으로서 TF팀에 참여해 직접 기획을 주도해볼 수 있다는 것 자체만으로도 많은 것을 배울 수 있었습니다! 행사 당일 참가자들이 분석을 수행하는 모습을 보면서 저도 함께 도전해 보고 싶다는 생각이 들었습니다. 다음 데이터톤에서는 운영진이 아닌 참가자로서 만나 뵙겠습니다!ㅎㅎ



# KHUDA DATATHON



## BIBIM

### THE 1ST KHUDA DATATHON

2024.11.09 SAT 09:00~21:00  
경희대학교 국제캠퍼스전자정보대학 211-1호

주제		비빔 제시한 데이터셋 중 두 가지 이상을 even하게 배워서 인사이트 도출
일시		11월 9일 (토) 09:00 ~ 21:00
장소		전자정보대학관 211-1호
접수기간		11월 4일 (월) 23:59 까지
접수방법		구글폼 (하단 QR코드)
참가자격		데이터를 활용해 문제 해결 능력을 기르고, 협업을 통해 데이터 분석 역량을 높이고 싶은 대학생 누구나 (대학생, 무대에 소속되지 않아도 가능)
참가비		인당 개인 부담금 10,000원 (KHUDA 지원금 15,000원)
참가혜택		점심/저녁 식사 및 다과 제공, 다른 데이터 분석 분야 학생들과의 교류와 네트워킹 기회
상금		최우수상(30만원) 1팀, 우수상(10만원) 2팀

타임 라인		09:00 ~ 09:10   참가자 등록 및 입장
		09:10 ~ 09:40   쿠다 데이터톤 오리엔테이션, 주제 공개
		09:40 ~ 10:00   현대모비스 후원사 홍보 세션
		10:00 ~ 13:00   오전 활동, 중간 과제 제출
		13:00   점심 식사 제공
		14:00 ~ 15:00   연사자 강연 : 코로나앱 개발자 이동훈님
		15:00 ~ 19:00   오후 활동, 최종 과제 제출
		19:00 ~ 19:30   저녁 식사 제공 및 우수팀 선정 투표 진행
		19:30 ~ 20:00   우수팀 발표
		20:00 ~ 21:00   심사 결과 발표 및 시상식, 사진 촬영

KHUDA

HYUNDAI  
MOBIS

활동사진  
(3매 이상)

← 홍보게시판  
경희대



데이터분석/AI 중앙동아리 KHUDA  
11/02 01:14

“데이터톤”이라 생각하니가 어려워보이는 거임

“대학생 총출동!! 두근두근 데이터분석 최강자 결정전 가을축제 스페셜!”이라고 생각하면 엄청 즐거워짐

👍 999+ 💬 999+ ☆ 999+

# 천하제일 데이터분석 최강자 결정전

주제	비밀 제1회 데이터톤 중 두 가지 이상을 even하게 비버서 인사이트 도움
일시	11월 9일 (토) 09:00 ~ 21:00
장소	전자정보대학관 211-1호
접수기간	11월 4일 (월) 23:59 까지
접수방법	구글폼 (하단 QR코드)
참가자격	데이터를 활용해 문제 해결 능력을 기르고, 협업을 통해 데이터 분석 역량을 높이고 싶은 대학생 누구나 (타학생, 무대에 소속되지 않아도 가능)
참가비	인당 개인 부담금 10,000원 (KHUDA 지원금 15,000원)
참가혜택	점심/저녁 식사 및 다과 제공, 다른 데이터 분석 분야 학생들과의 교류와 네트워킹 기회
상금	최우수상(30만원) 1팀, 우수상(10만원) 2팀

주최	경희대학교 중앙동아리 데이터분석/AI 동아리 KHUDA
후원	현대모비스

추가 정보는 @khu\_da.official 프로필의  
노션 [제1회 KHUDA 데이터톤] 홍보 페이지 참고

문의: 정유진 KHUDA 6기 회장 (010-6286-3181)  
KHUDA 인스타그램 DM @khu\_da.official













동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진



## 동아리 활동보고서

활 동 명	머신러닝 기초 세션		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 7. 24. ~ 2024. 8. 28. 매주 수요일 18:00~20:00	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다.</p> <p>2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>3. KHUDA 회칙 제4장 제14조 [정규 세션] 세션은 본 동아리의 핵심 교육 프로그램으로, 회원들에게 필요한 기초 지식을 제공하고 최신 동향을 소개한다. 방학 중에는 운영진이 주도하여 데이터분석, 머신러닝, 딥러닝 관련 기초 세션을 통해 활동 회원의 기본적 지식 보완 및 역량 향상을 지원한다.</p>		
활동내용 및 소감	<p>장소 : 공과대학 469호 및 B105</p> <p>활동내용 : 머신러닝 기초 세션은 KHUDA 신규 회원들이 필수적으로 이수해야 하는 학습 프로그램으로, 머신러닝의 기본 개념과 실습을 체계적으로 익힐 수 있도록 구성되었습니다.</p> <ul style="list-style-type: none"> <li>- 신규 회원들은 세션 전까지 정해진 학습 범위를 개별적으로 공부한 후 Github 블로그에 업로드합니다. 이후 세션 활동 시간에는 팀원들과 함께 실습과 과제를 수행하며 개념을 직접 적용해보는 기회를 가졌습니다.</li> <li>- 각 팀에는 운영진 학생이 팀장으로 배치되어, 신규 회원들의 학습 진행 상황을 확인하고 필요한 경우 개별적으로 도움을 줄 수 있도록 하였습니다. 이를 통해 신규 회원들이 어려움을 겪지 않고 안정적으로 학습을 이어갈 수 있도록 지원하였습니다.</li> <li>- 매주 5개의 팀이 돌아가며 학습 내용 및 심화 주제에 대한 발표를 준비하며, 이를 통해 머신러닝 이론을 더욱 깊이 있게 이해할 수 있도록 하였습니다.</li> <li>- 발표 팀은 해당 주치의 교재 내용을 바탕으로 실습 과제를 제작하여 준비해왔으며, 이를 통해 신규 회원들의 학습 수준을 점검하고 보완할 수 있는 시간을 가졌습니다.</li> <li>- 세션은 매주 수요일 18:00~20:00에 진행되었으며, 팀원 간 협업을 촉진하고 적극적인 참여를 유도하기 위해 다양한 토론 및 피드백이 이루어지도록 운영하였습니다.</li> </ul> <p><b>1) 머신러닝 기초 세션 활동 내용</b></p> <p>머신러닝 기초 세션의 구체적인 타임라인은 아래와 같습니다.</p>		

1) 나만의 작은 무기 발표 ( 18:00 ~ 18:15 )

- 매주 각 팀별 최소 1명 이상의 팀원이 **나만의 작은 무기 발표**를 진행합니다.

2) 학습 내용에 대한 발제 ( 18:15 ~ 18:35 )

- 발제를 맡은 인원은 이를 기반으로 **학습 내용에 대한 발제**를 진행합니다.
  - 각 주차별로 모든 팀원들은 정해진 학습 내용을 개인적으로 **깃허브**에 정리합니다.

3) 심화 주제 발표 ( 18:35 ~ 18:45 )

- 각 주차별로 발제를 맡은 팀에서 주차 학습 내용에 관한 심화 주제를 발표합니다.
- 주제는 **학습 내용과 연관된 어떠한 주제든 상관없으며, 학습한 내용이 실제로 어떻게 활용 및 응용 되는지**를 발표하는 것을 권장합니다.

4) 실습 및 과제 ( 18:45 ~ 19:30 )

- 실습 ) 학습한 내용을 기반으로 실습을 맡은 조는 다음과 같은 문제와 답변을 준비해야 합니다.
  - a. **코드가 어떤 의미를 가지는지 / 다른 동작을 수행하려면 어떻게 수정해야 하는지 등**
- 과제 ) 발표 조가 준비해온 **빈칸이 뚫린 코드**를 스스로 채워나가는 것을 말합니다. 과제는 **세션 다음날 자정**까지 제출합니다.

5) 이전 주차 학습 내용에 대한 질문 답변 ( 19:30 ~ 19:50 )

- 이전 주차 발표를 듣고, 팀 별로 정한 질문에 대한 답변을 준비해옵니다. 답변에 대해 토론하여 정답을 도출해냅니다.
- **세션장이 무작위로 선정하는 인원**은 팀의 질문과 답변에 대해 간단하게 발표합니다.

6) 질문 선정 ( 19:50 ~ 20:00 )

- 작성해온 질문들 중에서 **가장 중요한 질문 한 가지를 선정**합니다.

## 2) 벌점 제도 도입

모든 신규 회원이 머신러닝 기본 개념을 충실히 익힐 수 있도록 벌점 제도를 도입하였습니다. 운영진 학생들이 벌점 부과 사항을 꼼꼼히 확인하고 기록하여, 추가 보충 학습이 필요한 학생들을 파악하고 적절한 학습 지원을 제공할 수 있도록 관리하였습니다. 이를 통해 신규 회원들이 꾸준히 학습을 이어갈 수 있도록 유도하고, 동아리 내의 학습 문화를 더욱 체계적으로 정착시킬 수 있었습니다.

### ※ 벌점 부과 기준

- 무단 결석: -1점
- 무단 지각: -0.5점
- 과제 지각 제출: -0.5점
  - 사전 과제 및 코드 작성 과제 각각에 대해 별도로 부과
- 과제 미제출: -1.5점

누적 벌점이 -3점에 도달할 경우 제명될 수 있으며, -3점일 때 운영진과 면담이 필요합니다.

본인의 누적 벌점은 언제든지 운영진에게 문의하여 확인할 수 있습니다.

소감 :

- 채\*\* : 코딩에 대한 지식이 거의 전무에 가까운 상태에서 ML 기초세션부터 수강했는데, 배경 지식이 부재하다보니 세션의 내용을 온전히 배우기 어려웠던 면이 있었던 것 같습니다. 심화세션에서 데이터 비즈니스 세션을 선택해서 학기 중에 ml 세션에서 배웠던 내용을 다시 remind 할 수 있어서 좋았습니다. 또, 부교재로 통계 내용을 다뤘는데 응용통계학에서 배우는 내용과 매우 밀접해서 산공과 2-2에겐 최적의 심화 세션이었습니다. 그래서 무엇을 얻었냐면,, 전반적인 머신러닝과 통계적인 지식을 얻을 수 있었습니다.

활동사진  
(3매 이상)











동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

## 동아리 활동보고서

활동명	ML 기초 세션 토이프로젝트 컨퍼런스		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 4. 19:00~21:00	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 학습한 내용을 바탕으로 실용성 있는 프로젝트를 진행하여 실무 및 연구 역량을 증진한다.</p> <p>2. KHUDA 회칙 제4장 제15조 [컨퍼런스] 컨퍼런스는 회원들이 주도적으로 진행한 연구 결과를 공유하고 피드백을 받는 자리로, 팀 또는 개인 기반의 프로젝트 결과를 발표한다.</p>		
활동내용 및 소감	<p>장소 : 멀티미디어관 112호</p> <p>활동내용 :</p> <p>ML 기초 세션의 마무리 단계로, 신규 회원들이 학습한 내용을 실제 프로젝트에 적용해볼 수 있도록 ML 기초 세션 토이프로젝트 컨퍼런스를 개최하였습니다.</p> <p>이번 컨퍼런스는 KHUDA의 모토인 "함께 성장하며 한계를 뛰어넘는 여정"을 실천하는 자리로, 신규 회원들이 머신러닝 기초 세션에서 배운 지식을 현실 문제 해결에 적용하고, 프로젝트 경험을 쌓을 수 있도록 마련되었습니다.</p> <p>이번 토이프로젝트 컨퍼런스를 통해 신규 회원들은 머신러닝의 개념을 단순히 학습하는 것을 넘어, 실제 데이터에 적용하고 문제를 해결하는 과정을 경험할 수 있었습니다. 또한 교수님들의 피드백을 통해 프로젝트의 개선점을 찾고, 향후 발전 방향에 대한 인사이트를 얻는 기회가 되었습니다.</p> <p>또한, 본 행사는 KHUDA 신규 회원들이 팀워크를 경험하고, 머신러닝 실무 역량을 키울 수 있도록 지원하는 중요한 과정이었으며, 향후 더욱 발전된 프로젝트 기반 학습 문화를 정착시키는 데 기여할 것입니다.</p> <p>1) 토이프로젝트 개요 및 준비 과정</p> <p>토이프로젝트는 개인 또는 소규모 그룹이 여유 시간을 활용하여 단기간에 개발하는 프로젝트로, 신규 회원들이 머신러닝 이론을 실제 문제에 적용하는 기회를 제공하였습니다.</p> <ul style="list-style-type: none"> <li>- 프로젝트 준비 기간: 8/28 ~ 9/3</li> <li>- 참여 대상: KHUDA 신규 회원</li> <li>- 주제 선정: 학습한 머신러닝 개념을 활용하여, 현실적인 문제를 해결할 수 있는 프로젝트 기획하였습니다.</li> <li>- 멘토링 지원: 운영진 및 선배 회원들이 신규 회원들의 프로젝트 진행을 도와 피드백을 주고 받고, 프로젝트 발전 방향을 제시하였습니다.</li> </ul>		

## 2) 컨퍼런스 진행 및 평가 방식

컨퍼런스 당일, 소프트웨어융합학과 및 산업경영공학과 교수님들을 초청하여 프로젝트 발표 세션을 진행하였으며, 이를 통해 신규 회원들이 방학 동안 학습한 내용을 정리하고 공유하는 자리를 가졌습니다.

- 참석 교수님: 소프트웨어융합학과 박상근 교수님, 산업경영공학과 신정우 교수님, 산업경영공학과 최현홍 교수님
- 평가 방식: 회원 평가 점수 60% + 교수 평가 점수 40%
- 평가 기준은 운영진 내부 회의를 거쳐, 프로젝트의 완성도, 창의성, 기술성, 활용성, 전달성을 균형 있게 반영할 수 있도록 설정하였습니다.

### 세부 평가 기준

평가 항목	내용
완성도	프로젝트가 명확한 목표를 가지고 잘 마무리되었는지
창의성	아이디어와 접근 방식이 얼마나 참신했는지
기술성	사용된 머신러닝 기법의 난이도 및 적용 방식의 적절성
활용성	프로젝트 결과물이 실제로 활용될 가능성이 있는지
전달성	발표가 논리적으로 명확하게 전달되었는지

## 3) 팀별 주제

팀명	주제
1조	CV를 활용한 '무신사' 이미지 검색 모델
2조	급행열차 최적 정차역 제안
3조	KBO 경기 승부 예측 모델 개발
4조	서울특별시 프렌차이즈 카페 상권 추천
5조	폐의약품 무인 수거 기기의 최적 입지 선정 방안

## 4) 교수님 피드백

**A 교수님:** KHUDA에서 매 학기 머신러닝 기초 세션과 토이프로젝트 및 컨퍼런스를 운영하는 것으로 알고 있다. 이번 컨퍼런스에서도 학생들의 적극적인 참여와 수준 높은 프로젝트 결과물이 인상적이었다. 특히, 팀별 조언을 체계적으로 기록하고 이를 다음 학기 신규 회원들에게 전달한다면, 후속 기수의 프로젝트 완성도가 더욱 향상될 것으로 기대된다. 동아리가 지속적으로 발전할 수 있도록 이러한 피드백 시스템을 적극적으로 활용하면 좋겠다.

**B 교수님:** 짧은 준비 기간 동안 높은 완성도의 프로젝트를 준비하느라 고생 많았다. 단 일주일 만에 이 정도 수준의 결과물을 만들어냈다는 점이 매우 인상적이며, 학생들의 학습 속도와 문

제 해결 능력이 뛰어나다는 것을 확인할 수 있었다. 덕분에 제가 진행하는 수업에서도 과제 난이도를 조금 더 높여도 되겠다는 생각이 들었다. 앞으로도 실전 프로젝트 경험을 쌓으며 더욱 성장하길 바란다.

**C 교수님:** 프로젝트들이 머신러닝 기법을 활용하여 현실 문제를 해결하려는 시도를 했다는 점이 매우 긍정적이었다. 다만, 몇몇 팀은 모델 성능 개선이나 데이터 전처리 과정에서 추가적인 고민이 필요해 보였다. 다음 학기에는 프로젝트 기획 단계에서 데이터 수집 및 전처리에 대한 가이드라인을 조금 더 명확하게 제공하면, 더욱 발전된 결과물이 나올 수 있을 것이라 생각한다. 또한, 다양한 도메인의 데이터를 다루어보는 경험이 쌓이면 더욱 실용적인 프로젝트가 나올 것으로 기대된다.

**소감 :**

- **신\*\* :** 심화 세션과 프로젝트를 진행하면서, 열정적인 팀원들과 함께 협력하는 과정에서 팀플의 진정한 즐거움을 느꼈습니다. 모든 팀원이 각자 최선을 다하는 모습을 보면서 더욱 노력하였고, 이를 통해 기술적 역량뿐만 아니라 발표 능력, 문제 해결 능력 등 여러 가지를 배울 수 있었습니다. 특히, 서로의 강점을 살려 협업하고 목표를 이루는 과정에서 팀워크의 진정한 의미를 경험할 수 있었던 것 같습니다.



# KHUDA 6TH TOYPROJECT CONFERENCE



활동사진  
(3매 이상)

## WHEN & WHERE |

2024.09.04 (WED)  
19:00 ~ 21:00  
멀티미디어관 112호

## THEME |

- 1 CV를 활용한 '무신사' 이미지 검색 모델  
이수민 정유진 신수현 윤예준 이상엽 김정은 박지연 김민서 이준영
- 2 급행열차 최적 정차역 제안  
신경수 신헌석 김예람 송인혁 전지현 오윤진 박창희 장하늘
- 3 KBO 경기 승부 예측 모델 개발  
황종훈 정혜민 김형준 임주현 최진영 권도현 이준용
- 4 서울특별시 프렌차이즈 카페 상권 추천  
홍민혁 안형민 이재희 박정원 김혜란 채진기 문강민
- 5 폐의약품 무인 수거 기기의 최적 입지 선정 방안  
윤소은 허채은 장영주 김동혁 이 찬 장준필 이동진









동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

(정유진)

## 동아리 활동보고서

활 동 명	심화 세션 - 금융 트랙		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 18. ~ 2024. 11. 25. 매주 수요일 18:00~20:00	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>2. KHUDA 회칙 제4장 제14조 [정규 세션] 세션은 본 동아리의 핵심 교육 프로그램으로, 회원들에게 필요한 기초 지식을 제공하고 최신 동향을 소개한다. 학기 중에는 운영진이 주도하여 데이터분석, 머신러닝, 딥러닝 관련 심화 세션을 통해, 활동 회원의 역량 심화 및 실용적 지식을 제공한다.</p>		
활동내용 및 소감	<p>장소 : 공과대학 469호 (산업경영공학과 실습실)</p> <p>활동내용 :</p> <p>금융 서비스 산업은 방대한 데이터를 실시간으로 생성하며, 이를 해석하여 신속하고 정확한 의사 결정을 내리는 것이 필수적입니다. 특히, 머신러닝 기술은 금융 산업 전반에 걸쳐 큰 영향을 미치고 있지만, 알고리즘의 이론적 개념과 실제 구현 사이에는 상당한 괴리가 존재합니다. 이번 KHUDA 6기 금융 트랙에서는 이러한 문제를 해결하기 위해 금융 산업에서 중요한 머신러닝 기반 알고리즘을 구축하는 기술을 학습하고, 로보어드바이저 및 금융 데이터 분석 기법을 심층적으로 연구하는 것을 목표로 하였습니다.</p> <p>1) 주요 학습 내용 및 활동</p> <p>금융 트랙에서는 머신러닝과 딥러닝 기법을 활용하여 금융 데이터 분석을 수행하며, 실제 금융 전략에 적용할 수 있는 다양한 모델링 기법을 다루었습니다.</p> <p>기초적인 금융 도메인 지식에 관해서는 정규 심화세션에서 다루지 않고, 추가적인 금융 스터디를 진행하여 채워나갔습니다. 또한, 최종 컨퍼런스에서 발표할 프로젝트의 방법론과 적용법을 관련 서적에서 발췌하여 학습하였습니다.</p> <p>각 주차마다 팀원이 돌아가며 발제를 진행하였으며, <u>학습 내용을 심화 프로젝트에 어떻게 적용할 것인지에 대한 인사이트를 공유하는 데 집중</u>하였습니다. 해당 주차의 학습 내용을 Github 블로그에 작성하여 과제로 제출하도록 하였습니다.</p> <p>주요 학습 목표 및 내용</p>		



주요 학습 목표	내용
시계열 분석 및 알고리즘 구현	금융 데이터의 특성을 고려하여 시계열 기반 예측 모델을 개발
딥러닝을 활용한 금융 모델링	신경망 및 강화학습을 활용하여 금융 시장의 패턴 분석
백테스팅(Backtesting) 실습	개발된 모델의 성능을 평가하고 최적화하여 실제 적용 가능성을 검증하고자 함
포트폴리오 전략 최적화	다양한 금융 전략(멀티팩터, 이동평균선 활용 등)을 적용하여 투자 포트폴리오 최적화
금융 전략 구현의 어려움과 해결 방안	딥러닝, 강화학습, 자연어처리(NLP) 및 금융 전략을 통합하여 활용하는 과정에서 발생하는 난제들을 해결하고 함

## 2) 프로젝트 실제 적용 및 목표

금융 트랙에서는 수리·통계적 접근보다는 실무적 관점에서 사용자 특성(재무 상태, 투자 성향 등)과 거시경제 변수(GDP, 금리 등), 금융 지표(주가, 거래량, 이동평균선 등)를 활용하여 최적의 모델링을 수행하였습니다.

특히, 금융 모델을 활용하여 포트폴리오 조정 최적화 및 리밸런싱 자동화를 목표로 하였으며, 고객의 Financial Needs를 제한조건으로 설정한 최적화 문제 해결을 중점적으로 다루었습니다.

## 3) 사용 교재 및 참고 자료

금융 트랙에서는 다음과 같은 교재를 활용하여 학습을 진행하였습니다.

- 《금융 전략을 위한 머신러닝》
- 《파이썬으로 구현하는 로보어드바이저》
  - 3장: 평균-분산 전략 구현 및 시뮬레이션 분석
  - 7장: 멀티 팩터 전략
  - 8장: 딥러닝 예측을 통한 시장 모니터링
  - 9장: 고급 최적화 전략

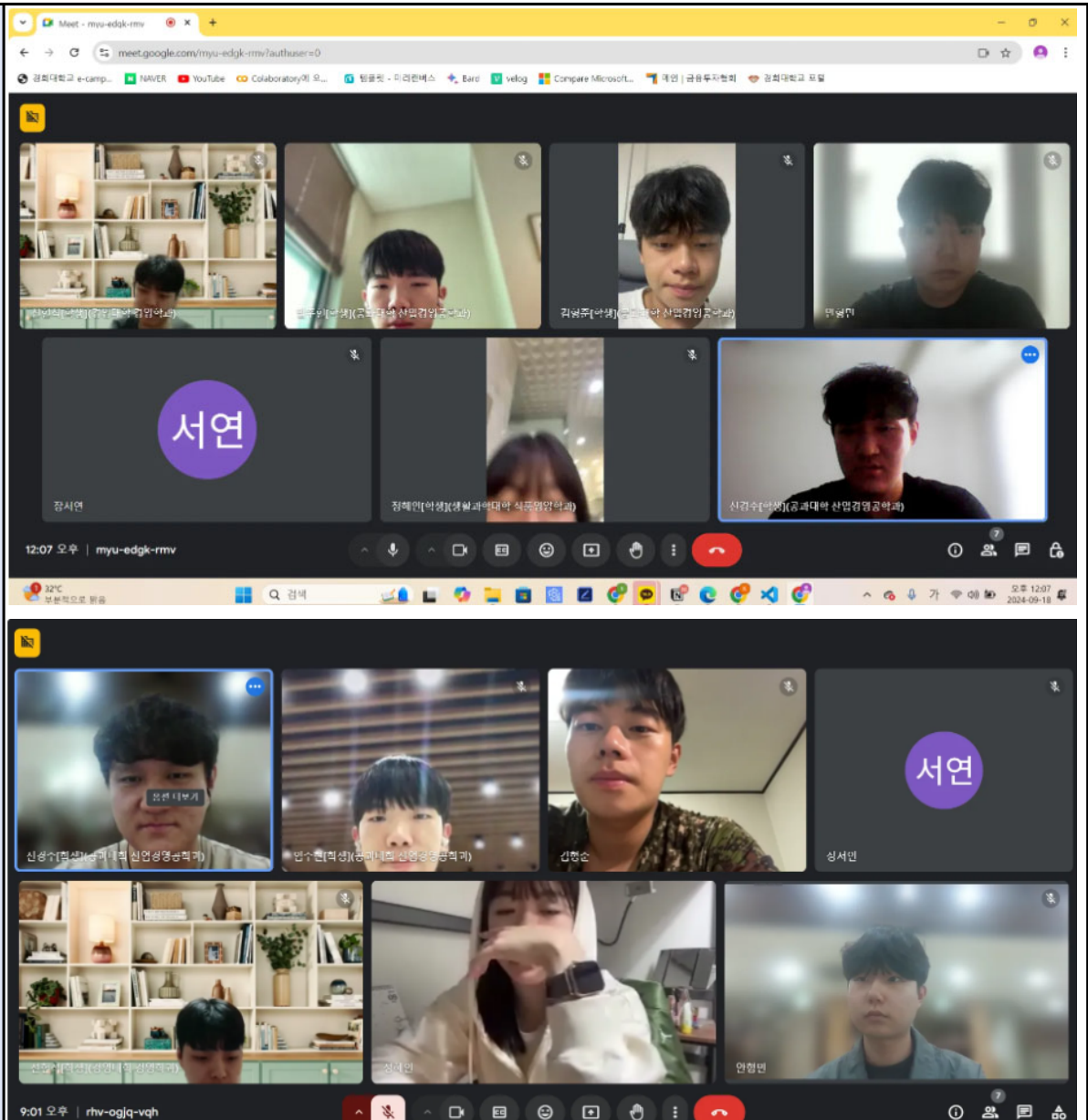
## 4) 트랙 커리큘럼

주차	날짜	학습 내용
OT	9/18	- 프로젝트 관련 회의 (스토리 보드) - 《금융 전략을 위한 머신러닝》 3장 인공지능경망
1	9/24	- 《금융 전략을 위한 머신러닝》 4장 지도학습 : 모델 및 개념 5장 지도 학습 : 회귀(시계열 모델)
2	10/1	- 《금융 전략을 위한 머신러닝》 6-3장 지도 학습 : 분류 (비트코인 거래 전략) 7장 비지도 학습 : 차원 축소
3	10/10	- 《금융 전략을 위한 머신러닝》 8장 비지도 학습 : 군집화. 9장 강화 학습
4	10/16	- 《금융 전략을 위한 머신러닝》 10장 자연어 처리 - 《파이썬으로 구현하는 로보어드바이저》 3장 평균-분산 전략 구현 및 시뮬레이션 분석
5	10/22	- 《파이썬으로 구현하는 로보어드바이저》 7장 멀티 팩터 전략
6	10/29	- 《파이썬으로 구현하는 로보어드바이저》 8장 딥러닝 예측을 통한 시장 모니터링 9장 고급 최적화 전략
중간고사 휴동		
9	11/6	심화 프로젝트 주제 논의
10	11/13	심화 프로젝트 진행
11	11/20	심화 프로젝트 진행
12	11/26	심화 프로젝트 발표 자료 준비 및 마무리

소감 :

- 정\*\* : 경제 지표나 주가나 모델링에 어떻게 활용되는지 막연한 느낌만 있었는데 실제로 구현하면서 데이터를 어떤 식으로 사용하게 되는지 확인해볼 수 있었습니다.

활동사진  
(3매 이상)







## Juhyun's codes

khu ie 21

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT



## Sun4-me

My path to becoming contributor from user

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT

# 금융트랙 2주차 정리

Posted Sep 24, 2024 • Updated Sep 24, 2024

By Ju hyun lim

14 min read

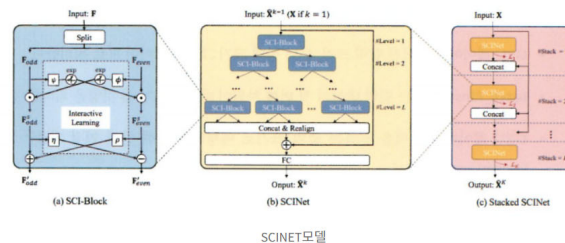
## 지도학습: 회귀(시계열 모델)

- 회귀 기반 지도 학습은 예측 모델로, 목표와 예측 변수간의 관계를 모델링하고 가능한 연속적 출력값을 예측하는 것 -> 금융에서 가장 많이 사용하는 머신러닝
- 자산 가격, 자산 수익 예측과 같은 투자 기회 <- 예측 변수의 속성, 다른 변수와의 관계를 이해하고 자산 수익을 견인하는 **중요한 요소**를 찾는것
- 뿐만 아니라 포트폴리오 관리, 보험 가격 책정, 상품 가격 책정, 헤징, 위험 관리 등 광범위한 재무 영역에 적용

## 사용되는 모델

- 시계열 모델: 과거에 일련의 데이터에서 무슨 일이 발생했는지 추론하고 앞으로 무슨 일이 발생할지 예측 하는 것
  - 고정성(추세나 계절성을 갖지 않는 것)을 없애고 자기상관을 반영하고자 AR(자기상관)(차분)MA(이동평균)모델을 주로 사용함
  - ARIMA 뿐만 아니라 ARIMAX(외생 변수가 포함되는, 시간 + a), SARIMA(계절성 요소를 모델링 하는 것이 목표), VARMA(다변수 확장-> 동시에 여러 변수 예측)을 사용하기도 함
  - 시계열과 딥러닝은 밀접한 연관이 있는데 LSTM(장단기 메모리)를 사용하면 예전의 시계열 정보

## SCINET



### Contents

- 딥러닝 예측 모델 구성
- 분석
- 딥러닝 모델
- RNN
- SCINET
- NLinear

SCINET 모델은 시간의 불규칙성을 고려하기 위해 입력 데이터의 짝수(even), 홀수(odd) 스텝을 구분하는 아주 독특한 방식을 제안했다.

### • (a) SCI-Block의 구조

SCI-Block의 그림을 보면 실제로 짝수 스텝과 홀수 스텝으로 분할되면서 상호 간의 네트워크를 연결해 학습하는 구조를 알 수 있다. 각 짝수, 홀수 스텝의 데이터는 컨볼루션 필터를 통해 시퀀스의 특성이 추출된다.

### • (b) 트리 구조

SCI-Block을 계층별로 쌓은 트리 구조를 나타낸다. 최종적으로 Zig-Zag 얼라인먼트를 통해 합친 후 완전히 연결돼 출력값을 도출한다. 이 구조를 (c) 그림과 같이 여러 개 쌓은 스택 형태로 구성할 수도 있다.

# 금융전략을 위한 머신러닝 CH.5 지도학습:회귀(시계열 모델)

신경수 · 2024년 9월 23일

Financial Machine Learning

팔로우

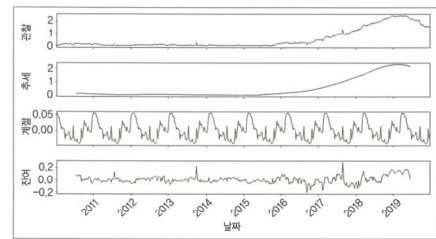


그림 5-1 시계열 구성요소

금융 산업에서는 상당히 많은 자산 모델링과 예측 문제가 시간 구성요소와 연속 출력의 추정과 관계된다.

시계열 분석 : 과거에 일련의 데이터에서 무슨 일이 발생했는지 추론하고 앞으로 무슨 일이 발생할지 예측하는 것

[시계열 모델]

1. 모수적임
2. 주로 예측 변수의 과거 데이터 사용
3. 예측을 위해 외생 변수를 사용할 수 있음

[지도 회귀 모델]

1. 비모수적임
2. 외생 변수를 예측 변수로 사용
3. 시간 지연 접근방식을 통해 예측 변수의 과거 데이터를 포함할 수 있음

-> 예측을 위해 예측 변수의 과거 데이터만이 아니라 외생 변수를 사용할 수 있다는 점에서 서로 유사함

+) 최종 출력 측면에서도 지도 회귀 및 시계열 모델을 변수의 가능한 결과의 연속 집합을 추정함

5.1 시계열 모델

- 5.1.1 시계열 연세
- 5.1.2 자기 상관과 고정성
- 5.1.3 기존 시계열 모델
- 5.1.4 시계열 모델링에 대한 딥러닝 접근방식
- 5.1.5 지도 학습 모델을 위한 시계열 데이터 수정
- 5.2 주제 1 : 주가 예측
- 5.3 주제 2 : 파생 상품 가격 책정
- 5.4 주제 3 : 투자자 위험 감수 및 로보 어드바이저
- 5.5 주제 4 : 수익률 곡선 예측

동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

## 동아리 활동보고서

활 동 명	심화 세션 – 자연어 처리 트랙														
동아리 분과	학술분과	동아리명	KHUDA												
활동시간	2024. 9. 11. ~ 2024. 11. 26. 매주 수요일 19:45 ~ 20:45	비고													
활동목적	1. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다. 2. KHUDA 회칙 제4장 제14조 [정규 세션] 세션은 본 동아리의 핵심 교육 프로그램으로, 회원들에게 필요한 기초 지식을 제공하고 최신 동향을 소개한다. 학기 중에는 운영진이 주도하여 데이터분석, 머신러닝, 딥러닝 관련 심화 세션을 통해, 활동 회원의 역량 심화 및 실용적 지식을 제공한다.														
활동내용 및 소감	<p>장소 : 전자정보대학 539호</p> <p>활동내용 :</p> <p>자연어 처리(NLP) 분야는 최근 몇 년간 <u>트랜스포머(Transformer)</u> 모델의 발전과 함께 비약적인 성장을 이루었습니다. 특히, GPT를 비롯한 대형 언어 모델이 NLP의 표준이 되면서, 기존의 RNN, LSTM, GRU 등의 모델은 점차 그 역할이 축소되었습니다.</p> <p>이번 KHUDA 6기 NLP 트랙에서는 최신 자연어 처리 기술을 학습하고, 트랜스포머 모델을 직접 구현하는 것을 목표로 하였습니다. 단순히 기존의 모델을 활용하는 것이 아니라, <u>PyTorch를 사용하여 핵심 요소(Self-Attention 등)를 직접 코드로 작성하며 모델의 작동 원리를 깊이 이해하는 것에 중점을 두었습니다.</u></p> <p><b>1) 주요 학습 내용 및 활동</b></p> <p>NLP 트랙에서는 트랜스포머 기반 언어 모델을 중심으로 이론과 실습을 병행하며 학습을 진행 하였습니다. Word2Vec, RNN, LSTM 등 기존 NLP 모델은 이미 ML 기초 세션/학과 교과 과정에서 학습하였으므로 다루지 않았습니다. 또한 사전에 정의된 토큰라이저를 활용하는 방법(Tokenizer) 또한 다루지 않았습니다.</p> <p><b>주요 학습 목표 및 내용</b></p> <table><tr><th>주요 학습 내용</th><th>설명</th></tr><tr><td>N-gram Language Models</td><td>전통적인 n-gram 기반 언어 모델의 개념과 한계를 학습</td></tr><tr><td>Transformers</td><td>트랜스포머 모델의 구조 및 Self-Attention 메커니즘 구현</td></tr><tr><td>트랜스포머 모델 직접 구현</td><td>PyTorch를 사용하여 트랜스포머 모델을 로우 레벨에서 구현</td></tr><tr><td>Hugging Face 활용</td><td>사전 학습된 트랜스포머 모델을 활용한 파인튜닝 기법 학습</td></tr><tr><td>모델 파라미터 튜닝</td><td>Pretraining, Fine-tuning 실습을 통한 성능 최적화</td></tr></table>			주요 학습 내용	설명	N-gram Language Models	전통적인 n-gram 기반 언어 모델의 개념과 한계를 학습	Transformers	트랜스포머 모델의 구조 및 Self-Attention 메커니즘 구현	트랜스포머 모델 직접 구현	PyTorch를 사용하여 트랜스포머 모델을 로우 레벨에서 구현	Hugging Face 활용	사전 학습된 트랜스포머 모델을 활용한 파인튜닝 기법 학습	모델 파라미터 튜닝	Pretraining, Fine-tuning 실습을 통한 성능 최적화
주요 학습 내용	설명														
N-gram Language Models	전통적인 n-gram 기반 언어 모델의 개념과 한계를 학습														
Transformers	트랜스포머 모델의 구조 및 Self-Attention 메커니즘 구현														
트랜스포머 모델 직접 구현	PyTorch를 사용하여 트랜스포머 모델을 로우 레벨에서 구현														
Hugging Face 활용	사전 학습된 트랜스포머 모델을 활용한 파인튜닝 기법 학습														
모델 파라미터 튜닝	Pretraining, Fine-tuning 실습을 통한 성능 최적화														

## 2) 진행 방식

NLP 트랙은 이론 학습, 실습, 세미나, 프로젝트의 네 가지 방식으로 구성되었습니다. 단순히 모델을 가져다 사용하는 것이 아니라, 트랜스포머 모델을 직접 구현하며 기초부터 심화까지 학습하는 과정을 목표로 하였습니다.

또한, 최신 논문과 연구를 분석하고 이를 실제 코드로 구현하는 과정을 통해 이론과 실전 적용의 균형을 맞추는 것을 중요하게 고려하였습니다.

### ① 이론 학습 및 실습 (매 주차 사전 과제)

**Andrej Karpathy의 유튜브 강의 Neural Networks: Zero to Hero를 기반으로 학습을 진행**하였습니다. 해당 강의는 인공지능 모델의 기초적인 개념부터 실제 구현까지 폭넓게 다루며, 특히로우 레벨에서 신경망을 직접 구현하는 방법을 강조하기 때문에 트랙 개설 목적을 달성하기에 최적의 강의라고 판단하였습니다.

이를 바탕으로 단순히 이론을 학습하는 것에서 그치지 않고, 실습을 병행하여 깊이 있는 이해를 도모하였습니다.

- **강의 수강 후 정리:** 학습한 내용을 블로그, 노션, GitHub 등 다양한 매체를 활용하여 정리했으며, 이를 통해 개념을 직접 정리하고 복습할 수 있도록 유도하였습니다.
- **실습 과제 수행:** 강의 내용을 바탕으로 PyTorch를 사용하여 모델을 직접 구현하고, 실험한 결과를 정리하여 GitHub에 공유하는 방식으로 진행했습니다.
- **트랙장이 제공하는 참고 자료 활용:** 특정 주제와 관련된 논문이나 문서를 추가적으로 학습하여 개념을 더욱 심화하였습니다.

### ② 세미나 진행 (세션 내 활동)

트랜스포머 모델을 심층적으로 이해하고 응용하기 위해 매주 '세미나' 형식으로 세션을 진행했습니다. 세미나를 진행하는 과정에서 트랙원들이 자유롭게 토론하고 발표하는 시간을 가졌습니다. 매 세션마다 정해진 발표 조가 진행을 담당하며, 팀원들과 함께 발표 자료를 준비하고 트랙원 모두와 공유하는 방식으로 진행했습니다.

- **Lecture Recap (~10분):** 매주 학습한 내용을 요약하여 발표합니다. 트랙원들이 주요 개념을 다시 한번 정리할 수 있도록 복습하는 시간을 가졌습니다.
- **자유 주제 발표 (20~30분):** 최신 NLP 연구나 응용 사례를 조사하여 발표하였습니다. Llama, RLHF, ChatGPT 등의 주제를 다루었으며, 트랙장이 매주 선정한 최신 논문을 참고하여 발표하는 것도 가능하도록 하였습니다.
- **Q&A 및 토론 (10~30분):** 발표 후에는 질의응답 시간을 가지면서 발표 내용에 대한 심도 있는 논의가 이루어지도록 하였습니다. 이를 통해 트랙원들이 보다 적극적으로 참여하고 서로의 이해도를 높일 수 있도록 유도하였습니다.

### ③ 프로젝트 진행 관련 준수 사항

트랜스포머 모델을 단순히 활용하는 데서 나아가, 모델을 최적화하고 성능을 개선하는 과정을 경험할 수 있도록 프로젝트 중심 학습을 진행하였습니다. 또한 프로젝트 진행에 있어서 단순히 **프롬프트 엔지니어링이나 단순 Retrieval-Augmented Generation(RAG) 기법을 사용하여 문제를 해결하는 방식은 지양**하였습니다. 최신 자연어 처리 기술을 능동적으로 적용하는 능력 함양과



함께 실무에서도 활용할 수 있는 프로젝트 경험을 쌓는 것이 트랙 개설의 목표였기 때문에, 보다 근본적인 모델 설계 및 튜닝을 포함한 접근을 장려하였습니다.

### 3) 사용 교재 및 참고 자료

NLP 트랙에서는 별도의 교재 없이 온라인 강의 및 최신 논문을 활용하여 학습을 진행하였습니다.

- 강의: Neural Networks: Zero to Hero (Andrej Karpathy, YouTube)
- 추가 논문 및 자료: 트랙장이 매주 최신 연구 및 논문을 선정하여 공유

### 4) 트랙 커리큘럼

#### 트랙 커리큘럼 ...

Aa 주차	≡ 학습 내용	≡ 진행 방식	≡ 강의	≡ Readings	≡ 발제자
1주차 (9/11)	Introduction to LLM	대면 이론	[1hr Talk] Intro to Large Language Models (youtube.com)		김재욱
🗣️ 2주차 (9/25)	Bigram Language Model	대면 실습	The spelled-out intro to language modeling: building makemore (youtube.com)		이동진 김동혁
🗣️ 3주차 (10/2)	N-gram Language Model	대면 실습	Building makemore Part 2: MLP (youtube.com)	bengio03a.dvi (jmlr.org)	오윤진 이상엽
🗣️ 4주차 (10/9)	Transformer part 1	이론	But what is a GPT? Visual intro to transformers   Chapter 5, Deep Learning (youtube.com)  주의력, 트랜스포머의 마음 시각화하기   6장, 딥 러닝 (youtube.com)	[1706.03762] Attention Is All You Need (arxiv.org)	
🗣️ 5주차 (10/16)	Transformer part 2	대면 실습	Let's build GPT: from scratch, in code, spelled out. (youtube.com)	[2005.14165] Language Models are Few-Shot Learners (arxiv.org)	김혜란 장영주
중간고사 휴동					
🗣️ 7주차 (10/30)	Transformer part 3	대면 실습	Let's build GPT: from scratch, in code, spelled out. (youtube.com)		권도현 박찬희
8주차 (11/6)	Fine-Tuning LLMs + 프로젝트	대면 실습	https://huggingface.co/learn/nlp-course/ko/chapter1/1  https://huggingface.co/learn/nlp-course/ko/chapter2/1  https://huggingface.co/learn/nlp-course/en/chapter3/1		김재욱
9주차 (11/13)	프로젝트 주제 선정	대면			
10주차 (11/20)	프로젝트 진행	대면			
11주차 (11/26)	프로젝트 발표 자료 준비 및 마무리	대면			

**소감 :**

- **이\*\*** : 심화 세션을 진행하면서 단순히 Transformer 기반의 LLM을 활용하는 것이 아닌, NLP의 시작인 Bigram부터 순차적으로 Transformer까지 모델에 대한 기초가 확립되었던 것 같습니다. Transformer 모델만 공부했었는데 왜 Transformer 모델을 사용하는지에 대한, 그리고 앞으로 Transformer 기반의 확장 모델의 MoE 등 많은 모델들의 이론적 학습과, 이를 실제로 레이어를 쌓아서 만들어보야 이해가 되겠다는 방향성을 제시해준 것 같아 좋았습니다. 프로젝트를 진행하면서 NLP라는 분야가 가지고 있는 고질적인 문제점 중 하나인 "데이터 확보"에 문제를 정말 정말 많이 겪었습니다. 뿐만아 **않은 것들을 하기 위해서는 실제 데이터 엔지니어링이 필요했고, 그 부분을 짧은 시간내에 해내기에는 쉽지 않았습니**다. 따라서, 공부와 실무의 괴리감을 많이 느꼈고, 앞으로 데이터 엔지니어링을 하면서 데이터를 핸들링 하면서 LLM 모델을 개발하기 보다는, 기존에 있던 데이터들을 활용해서 코드를 짜야겠다.

**활동사진  
(3매 이상)**





SangYub  
AI & ML study

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT



Chirpy  
A text-focused Jekyll theme

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT



## Transformer의 내부 구조 상세 설명

### 1. 입력 처리

- 토큰화 (Tokenization)
  - 입력 텍스트를 작은 단위(토큰)로 분할
  - 단어, 서브워드, 또는 문자 수준의 토큰화 가능
  - 예: "Hello, world!" → ["Hello", ",", "world", "!"]
- 임베딩 (Embedding)
  - 각 토큰을 고정된 길이의 벡터로 변환
  - 일반적으로 256에서 1024 차원의 벡터 사용
  - 학습 가능한 파라미터로, 모델 훈련 중 최적화됨

### 2. Transformer 블록의 주요 구성 요소

#### a. 어텐션 블록 (Attention Block)

- 목적: 각 토큰이 다른 모든 토큰과 어떻게 관련되는지 계산
- 작동 원리:
  - 각 토큰 벡터로부터 Query, Key, Value 벡터 생성
  - Query와 Key의 내적을 통해 어텐션 점수 계산
  - 점수를 정규화하고 Value에 가중치로 적용
- 멀티헤드 어텐션:
  - 여러 개의 어텐션 메커니즘을 병렬로 실행
  - 다양한 관점에서 토큰 간 관계 포착 가능

Home > NLP 심화발제

Search...

## NLP 심화발제

Posted Nov 6, 2024 • Updated Nov 6, 2024  
By your\_full\_name

16 min read

### Full Fine-Tuning

- Fine-tuning 중 하나로 사전 학습된 모델을 특정 작업에 맞게 재학습하여 최적화하는 과정
- 모델의 모든 가중치를 조정하는 방식
- 단점: 메모리와 계산 비용이 매우 많이 필요

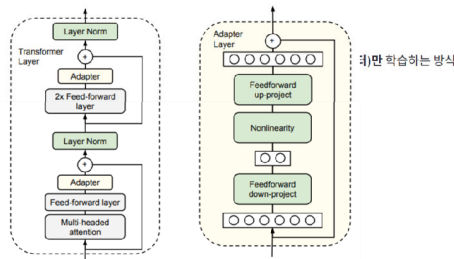
과정

### Parameter-Efficient Fine-Tuning(PET)

- 위의 비용문제를 해결하는 방식
- 전체 모델의 일부 파라미터만 조정하거나, 추가적인 모듈을 추가하여 효율적으로 fine-tuning 하는 방식

#### Adapter Tuning

- adapter는 대형 모델의 각 레이어에 추가되는 작은 모듈
- 기존 모델 가중치는 그대로 유지하면서 Adapter 모듈의 파라미터(추가파라미터)만 학습하는 방식
- 저차원에서 필요한 정보만 학습



이)만 학습하는 방식

- 오픈북: adapter, 원복: adapter를 transformer에 적용
- Transformer 레이어의 Feed-forward 층이나 레이어 정규화(Layer Norm) 직후에 삽입

#### Contents

- URL
- Predict, sample, repeat
- Transformer의 내부 구조 상세 설명
- 1. 입력 처리
- 2. Transformer 블록의 주요 구성 요소
- 3. 처리 과정
- 4. 최종 출력 생성
- 5. 주요 특징 및 개념
- 6. 스케일링 및 성능
- Chapter layout
- 단어 임베딩 (Word Embeddings)
- Unembedding
- Softmax with Temperature (온도 조절 ...
- Recap on embeddings
- Motivating examples
- The attention pattern (주의 집중 패턴)
- Masking
- Context size
- Values
- Counting parameters
- Cross-attention
- Multiple heads
- The output matrix
- Going deeper

#### Recently Updated

- NLP 심화발제
- NLP 5/6주차: transformer(2)
- NLP 4주차: transformer(1)
- NLP 3주차: mlp 모델로 이름 생성
- NLP 2주차: bigram 모델로 이름 생성

#### Contents

- Adapter Tuning
- 구조
- 성능
- LoRA
- QLoRA

## 과제 인증

Aa Name	🔗 1주차	🔗 2주차	🔗 3주차	🔗 4주차	🔗 5주차	🔗 7주차
권도현	<a href="https://github.com/riv...P-1.md">github.com/riv...P-1.md</a>	<a href="https://github.com/riv...P-2.md">github.com/riv...P-2.md</a>	<a href="https://riverblue72.github.io/">riverblue72.github.io/</a>	<a href="https://riverblue72.github.io/">riverblue72.github.io/</a>	<a href="https://riverblue72.github.io/">riverblue72.github.io/</a>	<a href="https://riverblue72.github.io/">riverblue72.github.io/</a>
김동혁	<a href="https://github.com/ddo...%AC.md">github.com/ddo...%AC.md</a>	<a href="https://ddonghyeok0326.github.io/jek...2.html">ddonghyeok0326.github.io/jek...2.html</a>	<a href="https://github.com/ddo...ipynb">github.com/ddo...ipynb</a>		<a href="https://github.com/ddo...ipynb">github.com/ddo...ipynb</a>	
김혜란	<a href="https://hyeran1216.github.io/202...-week1">hyeran1216.github.io/202...-week1</a>	<a href="https://hyeran1216.github.io/202...-week2">hyeran1216.github.io/202...-week2</a>	<a href="https://github.com/hye...ek3.md">github.com/hye...ek3.md</a>	<a href="https://hyeran1216.github.io/">hyeran1216.github.io/</a>	<a href="https://hyeran1216.github.io/202...-week5">hyeran1216.github.io/202...-week5</a>	<a href="https://hyeran1216.github.io/202...-week5-7">hyeran1216.github.io/202...-week5-7</a>
박찬희	<a href="https://notion.so/1we...d271a6">notion.so/1we...d271a6</a>	<a href="https://github.com/cha...ipynb">github.com/cha...ipynb</a>	<a href="https://github.com/cha...SSION">github.com/cha...SSION</a>	<a href="https://notion.so/4we...9a941d">notion.so/4we...9a941d</a>	<a href="https://github.com/cha...ipynb">github.com/cha...ipynb</a>	<a href="https://notion.so/7we...e6baa3">notion.so/7we...e6baa3</a>
오윤진	<a href="https://soft-mountain-353.notion.site/e/1-a...?pvs=4">soft-mountain-353.notion.site/e/1-a...?pvs=4</a>	<a href="https://soft-mountain-353.notion.site/e/2-1...?pvs=4">soft-mountain-353.notion.site/e/2-1...?pvs=4</a>	<a href="https://velog.io/@ko...A6%AC/">velog.io/@ko...A6%AC/</a>	<a href="https://velog.io/@ko...ers-정리">velog.io/@ko...ers-정리</a>	<a href="https://velog.io/@ko...lusion">velog.io/@ko...lusion</a>	<a href="https://velog.io/@ko...lusion">velog.io/@ko...lusion</a>
이상엽	<a href="https://khusangyub.github.io/pos...LP1주차/">khusangyub.github.io/pos...LP1주차/</a>	<a href="https://khusangyub.github.io/pos...B0%A8/">khusangyub.github.io/pos...B0%A8/</a>	<a href="https://github.com/KHU...%A8.md">github.com/KHU...%A8.md</a>	<a href="https://khusangyub.github.io/pos...B0%A8/">khusangyub.github.io/pos...B0%A8/</a>	<a href="https://khusangyub.github.io/pos...B0%A8/">khusangyub.github.io/pos...B0%A8/</a>	<a href="https://khusangyub.github.io/pos...B0%A8/">khusangyub.github.io/pos...B0%A8/</a>
이동진	<a href="https://actiondongjin.github.io/posts/ML1/">actiondongjin.github.io/posts/ML1/</a>	<a href="https://actiondongjin.github.io/posts/nlp1/">actiondongjin.github.io/posts/nlp1/</a>	<a href="https://actiondongjin.github.io/posts/nlp2/">actiondongjin.github.io/posts/nlp2/</a>	<a href="https://actiondongjin.github.io/posts/nlp4/">actiondongjin.github.io/posts/nlp4/</a>	<a href="https://actiondongjin.github.io/posts/nlp5/">actiondongjin.github.io/posts/nlp5/</a>	
장영주	<a href="https://0ju428.notion.site/1-l...?pvs=4">0ju428.notion.site/1-l...?pvs=4</a>	<a href="https://0ju428.notion.site/2-B...?pvs=4">0ju428.notion.site/2-B...?pvs=4</a>	<a href="https://github.com/you...ipynb">github.com/you...ipynb</a>	<a href="https://0ju428.notion.site/4-T...?pvs=4">0ju428.notion.site/4-T...?pvs=4</a>	<a href="https://github.com/you...ipynb">github.com/you...ipynb</a>	<a href="https://github.com/you...ipynb">github.com/you...ipynb</a>

동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진



## 동아리 활동보고서

활동명	심화 세션 - 데이터 비즈니스 & 추천시스템 트랙		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 18. ~ 2024. 10. 26.	비고	



	매주 수요일 18:00 ~ 20:00		
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>2. KHUDA 회칙 제4장 제14조 [정규 세션] 세션은 본 동아리의 핵심 교육 프로그램으로, 회원들에게 필요한 기초 지식을 제공하고 최신 동향을 소개한다. 학기 중에는 운영진이 주도하여 데이터분석, 머신러닝, 딥러닝 관련 심화 세션을 통해, 활동 회원의 역량 심화 및 실용적 지식을 제공한다.</p>		
활동내용 및 소감	<p>장소 : 중앙도서관 스터디룸 및 공과대학 편의점</p> <p>활동내용 :</p> <p>오늘날 데이터는 비즈니스의 핵심 자원으로 자리 잡았으며 이를 효과적으로 활용하는 것이 기업의 경쟁력을 좌우하는 중요한 요소로 자리 잡았습니다. 이에 따라 데이터 비즈니스 분야 또한 데이터를 통해 비즈니스 가치를 창출하고, 보다 정교한 의사결정을 지원하는 방향으로 발전하고 있습니다.</p> <p>데이터 비즈니스 &amp; 추천 시스템 트랙은 데이터 기반의 비즈니스 의사결정을 효과적으로 수행하는 방법과 추천 시스템의 원리를 학습하는 과정으로 구성하였습니다. 단순한 데이터 분석 기법을 익히는 것에 그치지 않고, 실제 데이터셋을 활용하여 '실무 환경에서 학습 내용을 어떻게 적재 적소에 효과적으로 적용할 수 있을지'를 고민하며 커리큘럼을 구성했습니다.</p> <p>본 트랙에서는 데이터 분석의 핵심이 되는 통계 개념과 머신러닝 기법을 학습하고, 이를 바탕으로 실제 비즈니스 문제 해결에 적용하는 경험을 쌓는 것을 목표로 삼았습니다. 이를 위해 본 교재 및 부교재, 온라인 강의 등 다양한 학습 자료를 활용하여 체계적인 학습을 진행했습니다. 더불어 프로젝트를 통해 실제 비즈니스 문제를 해결하는 과정을 경험하며, 데이터 활용 역량을 한층 더 높일 수 있도록 노력했습니다.</p> <p>(※ 신청 인원이 많아 2개의 팀으로 나누어 진행)</p> <p><b>1) 주요 학습 내용</b></p> <p>실무 역량을 강화하기 위해 '비즈니스 문제를 다루고 있는' 본교재를 중심으로 학습했습니다. 그리고 데이터 분석에 필수적인 통계 지식을 함께 함양하고자 통계 부교재를 추가적으로 활용하였습니다. 추천 시스템 학습은 마지막 3개 주차 동안 트랙원들이 돌아가면서 각자 한 가지 추천 시스템 알고리즘을 선정하여 발표하는 방식으로 진행했습니다. <b>1인 1 알고리즘 발제를 통해 다양한 알고리즘의 원리를 깊이 있게 이해</b>하고, 서로의 학습 내용을 공유하며 실무 적용 역량을 강화하는 기회가 되었습니다. 본교재의 핵심 내용은 다음과 같습니다.</p> <p><b>주요 학습 목표 및 내용</b></p>		

주요 학습 내용	설명
데이터 통계량 계산	데이터의 기본적인 통계 지표를 계산하여 분포와 특성 분석
데이터 가공 방법	결측값 처리, 이상치 제거, 데이터 변환 등의 데이터 전처리 기법 실습. 데이터 특성에 맞는 정제 방법 학습
데이터셋 결합	서로 다른 데이터셋을 결합하고 이를 활용한 분석 진행
클러스터링 결과 분석 및 시각화	군집화 기법을 적용하고, 결과를 분석 및 시각화하여 직접 인사이트를 도출하기
예측 모델 구현	Decision Tree를 이용한 예측 모델 구축 및 성능 평가
물류 최적 경로 문제 해결	비용 함수와 제약 조건을 고려한 최적 경로 탐색 알고리즘 학습
수치 시뮬레이션	시뮬레이션 기법을 적용하여 모델의 파라미터를 최적화하고 미래 예측 수행

## 2) 진행 방식

본 트랙은 **이론 학습과 실습을 병행하는 방식**으로 운영했습니다. 본교재와 부교재를 활용하여 매주 정해진 학습 범위를 학습하고, 이를 바탕으로 실습을 수행하는 형태로 진행했습니다.

트랙원들은 세션 전까지 정해진 교재 범위에 대해 개별적으로 학습한 내용을 정리하여 Github 블로그에 업로드하여 과제를 제출했습니다. 교재에 있는 실습 과정을 통해 학습한 내용을 실제 코드로 구현하고, 그 결과를 GitHub에 공유하도록 했습니다.

### ① 학습 운영

본교재는 실습 중심의 구성으로 이루어져 있어, 학습 중 이론적 개념이 부족한 경우 추가적인 학습이 필요한 경우가 많았습니다. 이에 따라 트랙원들은 필요한 이론을 개별적으로 보완하고, 학습 내용을 블로그에 정리하여 공유하는 방식을 활용했습니다.

또한 주어진 학습 범위를 보다 깊이 있게 이해하기 위해 주차별 과제(실습 코드 제작)를 추가적으로 운영했습니다.

- 매주 정해진 커리큘럼을 기반으로 학습 진행
- 학습 내용(본교재&부교재)을 화요일 24시까지 블로그에 정리 후 노선에 URL 업로드

### ② 세션 진행

정기 세션에서는 이론 발표와 실습을 병행했습니다. 이론 발제는 최소화하여 진행하고, 실습 과정에서 막힌 어려움을 트랙원과 공유하는 것을 가장 중요하게 생각했습니다. 이러한 일련의 과정을 통해 학습 내용을 체계적으로 정리하고, 본인의 것으로 만들어 실무 역량을 강화할 수 있도록 했습니다. 구체적인 세션의 진행 방식은 다음과 같습니다.

- 1. 본교재 & 부교재 발제:** 발제자가 해당 주차의 주요 개념 및 학습 내용을 발표합니다.
- 2. 퀴즈 풀이:** 매주 트랙장은 학습한 내용을 점검할 수 있도록 퀴즈를 준비했습니다.
- 3. Q&A 및 토론:** 교재 학습 중 발생한 질문 사항에 대해 깊게 논의하는 시간을 가졌습니다.
- 4. 실습 진행:** 준비된 실습 파일의 빈칸 코드를 채우며 빠르게 복습 진행

(실습 담당 트랙원이 해당 단원과 관련된 데이터셋을 직접 찾아, 교재에서 배운 내용을 적용할 수 있도록 실습 파일을 제작)

3) 사용 교재 및 참고 자료

- 본교재: 《파이썬 데이터 분석 실무 테크닉 100》(실무 기반 데이터 분석 및 머신러닝 실습)
- 부교재: 《통계 101×데이터 분석》(추론 통계, 가설 검정, 베이지안 통계 등 핵심 개념 학습)
- 추가 강의: 인프런 Python을 이용한 개인화 추천시스템 강의

4) 트랙 커리큘럼

트랙 커리큘럼

Aa 주차	≡ 학습 내용	≡ 진행 방식
 1주차 (9.18)	본교재 1장, 2장 / 부교재 1장, 2장, 3장	<span>대면</span> <span>비대면</span>
 2주차 (9.25)	본교재 3장, 4장 / 부교재 4장, 5장	<span>대면</span>
 3주차 (10.2)	본교재 5장, 6장 / 부교재 6장, 7장	<span>대면</span>
4주차(10/9)	본교재 7장, 8장 / 부교재 8장, 9장	<span>대면</span>
중간고사 휴동		
7주차(10.30)	본교재 9장, 10장 / 부교재 10장, 11장	<span>대면</span>
8주차(11.6)	추천시스템 / 부교재 12장, 13장	<span>대면</span>
9주차(11.13)	추천시스템 발제 심화 프로젝트 주제 논의	<span>대면</span>
10주차 (11.20)	추천시스템 발제 심화 프로젝트 주제 선정 및 진행	<span>대면</span>
11주차 (11.27)	추천시스템 발제 심화 프로젝트 PPT 제작 및 마무리	<span>대면</span>

소감 :

- 전\*\* : 통계학에 대한 개념만 있었는데 심화세션에서 상황에 맞게 적용하는 방법을 배우고 그것을 프로젝트에서 사용할 수 있었다. 교재 마지막 부분에 **물류 데이터를 다루는 부분**이 있었는데 프로젝트를 하는데에 도움이 되었고, 데이터 분석 외에도 그래프에 대한 기본적인 지식을 얻을 수 있어서 좋았다.

활동사진  
(3매 이상)









## 우당탕탕 데이터분석 성장기

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT

### 데이터 분석: 파이썬 데이터 분석 실무 테크닉 1008장

Posted Oct 7, 2024

by 김재현

데이터 분석: 파이썬 데이터 분석 실무 테크닉 1008장

Contents >

14 min read



#### 8장: 수치 시뮬레이션으로 소비자의 행동을 예측하는 테크닉 10

미래 예측을 위한 선택지를 넓히는 수업으로 수치 시뮬레이션!

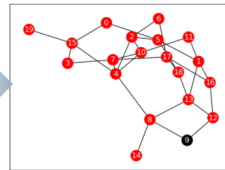
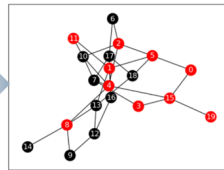
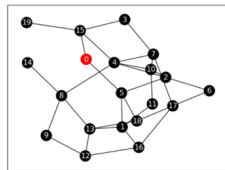
소비자 행동이 입소문과 같은 전파로 어떻게 변화하는지를 인간관계 네트워크 구조를 이용한 시뮬레이션

Problem) 우리 회사 제품의 판매를 예측할 수 있나? 우리 제품은 SNS를 통해 입소문을 타고 퍼지는 특성이 있어요.. 재구매 고객은 SNS로 연결돼 있어 파악이 가능하고... 이것을 이용해 앞으로의 매출 예측 가능??

```
1 # 2025
2
3 # 시계열 그래프 그리기
4 for i in range(len(list_timeSeries)):
5     list_timeSeries_num.append(sum(list_timeSeries[i]))
6
7 plt.plot(list_timeSeries_num)
8 plt.show()
```

```
1 # 2025
2
3 # 12
4 plt.plot(list_timeSeries_num, color="r", marker="o", linestyle="solid", markersize=10, markershape="circle")
5 plt.show()
```

```
1 # 2025
2
3 # 36
4 plt.plot(list_timeSeries_num, color="r", marker="o", linestyle="solid", markersize=10, markershape="circle")
5 plt.show()
```



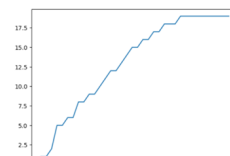
12개월까지는 원만한 전파였지만

36개월이 지나면 거의 전원에게 전파되는 모습!!

#### 3. 입소문 수의 시계열 변화를 그래프화해 보자

```
1 # 시계열 그래프 그리기
2
3 list_timeSeries_num = []
4 for i in range(len(list_timeSeries)):
5     list_timeSeries_num.append(sum(list_timeSeries[i]))
6
7 plt.plot(list_timeSeries_num)
8 plt.show()
```

각 시간 단계별 활성화된 노드의 수를 list\_timeSeries\_num 리스트에 추가시키면서 확인!



단계적으로 입소문이 나서 전원에게 전파되는 모습!!

Search...

#### Recently Updated

- 데이터 분석: R 스타디 4~5장
- 데이터 분석: R 스타디 1~3장
- 데이터 분석: lavaan, 구조방정식모형(SEM)
- 데이터 분석: 12장 통계 기반 분석 방법론
- 데이터 분석: 13장 선형회귀분석과 Elastic...

#### Contents

1. 인간관계 네트워크를 가시화해 보자
2. 입소문에 의한 정보 전파 모습을 가시...
3. 입소문 수의 시계열 변화를 그래프화...
4. 회원 수의 시계열 변화를 시뮬레이션...
5. 파라미터 전체를 '상관관계'를 보면서 ...
6. 실제 데이터를 불러와 보자
7. 링크 수의 분포를 가시화해 보자
8. 시뮬레이션을 위해 실제 데이터로부...
9. 실제 데이터와 시뮬레이션을 비교하자
10. 시뮬레이션으로 미래 매출 예측해 보자



## Jgzkan's notebook

Welcome to my archive!

- HOME
- CATEGORIES
- TAGS
- ARCHIVES
- ABOUT



### Ridge와 Lasso 그리고 Elastic Net

- Ridge(L2-norm): 전체 변수를 모두 유지하면서 각 변수의 계수 크기를 조정함. 예측에 영향을 거의 미치지 않는 변수는 0에 가까운 가중치를 주어 독립변수들의 영향력 조정. -> 계수 정규화(Regularization), 다중공선성을 방지하면서 모델 설명력 최대화 가능
- Lasso(L1-norm): Ridge와 유사하지만, 중요한 몇 개의 변수만 선택하고 나머지 변수들은 계수를 0으로 주어 영향력을 아예 없앤다. -> 모델 단순화, 해석 용이

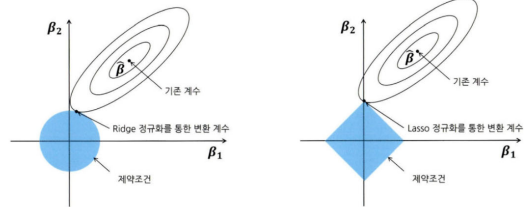


그림 13.5 Ridge와 Lasso의 산술 개념

- Elastic Net: Ridge와 Lasso의 최적화 지점이 다르기에 두 정규화 항을 결합하여 절충한 모델. Ridge는 변환된 계수가 0이 될 수 없지만 Lasso는 0이 될 수 있다는 특성을 결합한 것. Ridge와 Lasso의 혼합비율( $\alpha$ )을 조절하여 Elastic Net 모델의 성능을 최적화. ( $\alpha=0$ : Ridge,  $\alpha=1$ : Lasso)

표 13.2 다중 회귀분석 모형 결과

	DF	Sum of Squares	Mean Squares	F value	Pr > F
Model	2	7215.63	3607.82	27.23	( $\alpha$ ) <.0001

#### Contents

- 12. 통계 기반 분석 방법론
- 13. 머신러닝 분석 방법론
  - 13.1 선형 회귀 분석과 Elastic Net(에...
    - Ridge와 Lasso 그리고 Elastic Net
  - 부교재
  - 10장 인과와 상관
  - 11장 베이즈 통계





과제 제출

표

사전과제 제출

Aa 이름	과제 제출 (본교재)	과제 제출 (부교재)	Last edited time	실습파일 제출
최진영	github.com/cho...ipynb	choiwlsgithub.io/dat...3.html	2024년 10월 2일 오후 7:50	https://colab.rese...
윤예준	yejunyoon1102.github.io/cat...nique/	yejunyoon1102.github.io/pos...a6to7/	2024년 10월 2일 오후 7:49	3주차_실습.ipynb
송인혁	thddlsgr0105.github.io/c at...ssion/	thddlsgr0105.github.io/c at...ssion/	2024년 10월 2일 오후 7:51	3주차_실습.ipynb
박지연	https://cozyeon.github.io/posts/databusiness5/ https://cozyeon.github.io/posts/databusiness6/	cozyeon.github.io/pos...tics3/	2024년 10월 2일 오후 8:00	https://colab.rese...
전지현	github.com/jhy...rkdown	github.com/jhy...rkdown	2024년 10월 2일 오후 7:57	3주차_실습.ipynb
채진기	jgzkan.github.io/pos...%B0%A8	jgzkan.github.io/pos...%9E%AC	2024년 10월 2일 오후 8:08	https://colab.rese...
박정현	github.com/Jeo...3_3.md	github.com/Jeo...6_7.md	2024년 10월 2일 오후 8:07	https://github.co...

세션 발제, 실습자료

표

발제, 실습자료

Aa Name	Tags	자료
윤예준	발제	yejunyoon1102.github.io/cat...nique/
윤예준	발제	unequaled-point-f08.notion.site/113...0f146a
최진영	심화발제	choiwlsgithub.io/in-...1.html
채진기	발제	jgzkan.github.io/pos...A0%9C/



과제 제출

☞ 표

사전과제 제출

Aa 이름	🔗 과제 제출 (본교재)	🔗 과제 제출 (부교재)	🕒 Last edited time	🔗 실습
최소영	github.com/git....ipynb	velog.io/@so...%9E%A5	2024년 10월 3일 오전 12:57	github.com/git....ipynb
이준영	junyeong-khu.github.io/파이썬...크닉-10/	junyeong-khu.github.io/통계...-가설검정/	2024년 10월 3일 오전 11:33	github.com/jun....ipynb
최예지	github.com/Les...ession	lessie0501.tistory.com/	2024년 10월 3일 오전 2:38	github.com/Les....ipynb
허채은	chereunii.github.io/cat...84%9D/	chereunii.github.io/cat...84101/	2024년 10월 3일 오전 12:00	github.com/che....ipynb
이제희	pinkkj.github.io/cat...9-100/	pinkkj.github.io/cat...84%9D/	2024년 10월 3일 오후 2:57	github.com/pin....ipynb
김정은	github.com/sta...%93%9C	github.com/sta...rkdwn	2024년 10월 3일 오전 12:34	github.com/sta....ipynb
장하늘	sky-coding-24.tistory.com/15	sky-coding-24.tistory.com/16	2024년 10월 3일 오전 10:11	github.com/sns....ipynb

+ 새 페이지

세션 발제, 실습자료

☞ 표

발제, 실습자료


Aa Name	🏷 Tags	🔗 자료
이준영	본교재	junyeong-khu.github.io/categories/
최예지	부교재	lessie0501.tistory.com/
김채영	실습	colab.research.google.com/dri...haring

동아리활동과 관련하여 상기 결과보고서를 확인합니다.  
※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

## 동아리 활동보고서

활 동 명	심화 세션 - 컴퓨터 비전 트랙		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 9. ~ 2024. 11. 26. 매주 월요일 17:00 ~ 18:40	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>2. KHUDA 회칙 제4장 제14조 [정규 세션] 세션은 본 동아리의 핵심 교육 프로그램으로, 회원들에게 필요한 기초 지식을 제공하고 최신 동향을 소개한다. 학기 중에는 운영진이 주도하여 데이터분석, 머신러닝, 딥러닝 관련 심화 세션을 통해, 활동 회원의 역량 심화 및 실용적 지식을 제공한다.</p>		
활동내용 및 소감	<p>장소 : KHUDA 동아리방</p> <p>활동내용 :</p> <p>Computer Vision이라 하면, 아무래도 Face recognition이나 Object recognition 등이 가장 먼저 떠오르시리라 생각합니다. 그 외에도 Computer Vision을 통해서 재미있는 Task들을 해결할 수 있는데요, 그 예시로 Color style transfer를 들고자 합니다.</p> <p>Color style transfer란, Input Image(=Content Image)가 Style Image의 Color style을 따르게 style transfer를 진행하는 Task입니다. 아래는 논문 'Neural Preset for Color Style Transfer'에 첨부된 이미지입니다.</p> <div style="text-align: center;">  <p>Input Image    Style Image    Ours    CT    WCT<sup>2</sup>    PhotoNAS    PhotoWCT<sup>2</sup>    Deep Preset</p> </div> <p>Zhanghan Ke, et al. "Neural Preset for Color Style Transfer(2023)"</p> <p>이와 같이 Computer Vision(이하 CV)은 이미지 인식(Recognition), 객체 탐지(Detection)뿐만 아니라 다양한 분야에서 활용될 수 있는 기술입니다. CV 트랙에서는 이러한 가능성을 탐구하며, 관련 기술을 심층적으로 학습하고 코드 리뷰를 통해 실무적인 역량을 기르는 것을 목표로 했습니다.</p> <p>트랙 활동은 Deep Learning Specialization 강의(DeepLearning.AI, Coursera)를 기반으로 진행했습니다. 트랙을 통해 논문을 읽고 정리하는 능력을 기를 수 있도록 구성했습니다. 또한 Pytorch 프레임워크를 활용한 코드 리뷰를 진행하며 실제 모델 구현 및 개선 과정을 경험했습니다.</p>		

## 1) 학습 교재 및 참고 자료

### ① 학습 교재

DeepLearning.AI의 Coursera 강의인 Deep Learning Specialization을 학습 자료로 활용했습니다. 특히 다음 두 개의 Lesson을 중심으로 학습을 진행했습니다.

- Lesson 4: Convolutional Neural Network
- Lesson 5: Sequence Model

해당 강의에서는 기본적인 CNN 개념부터 Sequence Model을 활용한 Vision 태스크까지 다루고 있으며, 이를 통해 CV 트랙원들은 실무에서 적용할 수 있는 Computer Vision 기법을 익혔습니다.

### ② 참고 자료

- **Stanford University CS231n:** Computer Vision을 보다 심층적으로 학습하기 위해 Stanford University의 CS231n 강의를 추가적으로 참고했습니다.
- **트랙장이 진행한 프로젝트 코드:** 트랙장이 2023년 소프트웨어융합개론 과제를 시작으로, 2024년 6월까지 보완한 프로젝트 코드를 코드 리뷰에 활용했습니다. 교수님의 피드백을 반영한 코드이므로, 코드 리뷰 시 트랙원들이 분석한 문제점과 교수님의 피드백을 비교하며 학습할 수 있었습니다.

## 2) 진행 방식

### ① 사전 과제

트랙원들은 매주 주어진 학습 범위에 맞춰 사전 과제를 수행했습니다.

- **트랙원 전원**
  1. 각 주의 학습 진도에 맞춰 Deep Learning Specialization 강의를 수강했습니다.
  2. 학습 내용을 정리하여 노선에 제출했습니다.
  3. 해당 주차의 코드(트랙장의 프로젝트 코드)를 읽고, 코드 구조 및 문제점 분석 후 노선에 제출했습니다.
- **발제자**
  1. 강의 내용을 심화 학습하거나 관련 트렌드를 조사하여 발제를 준비했습니다.
  2. 해당 강의 내용과 관련된 쪽지 시험 문제를 출제했습니다.

### ② 세션 진행

정기 세션은 **발제 발표, 쪽지 시험, 코드 리뷰**로 구성되었으며, 학습한 내용을 체계적으로 정리하고 실무 적용 역량을 높이는 데 초점을 맞췄습니다.

- 40분: 발제자가 해당 주차 강의 내용 발표
- 20분: 쪽지 시험 진행
- 40분: 코드 리뷰 진행

### 3) 코드 리뷰 방식

코드 리뷰는 트랙장이 진행한 프로젝트 코드를 기반으로 진행되었으며, **Pytorch 프레임워크**를 활용했습니다. 다만 **Pytorch 자체를 학습하는 것이 아니라, 모델 구현 과정에서 발생한 문제점과 개선 방법을 중심으로 리뷰**가 이루어졌습니다.

트랙장이 진행한 프로젝트 코드는 총 세 개의 버전(Version 1, 2, 3)으로 구성되었으며, 각 버전의 핵심 내용은 다음과 같습니다.

Version	Backbone	Performance
1	VGG16	200 epoch 이상 학습을 진행했음에도 불구하고 accuracy 개선 없음
2	Inception	50 epoch 이상 학습했으나 accuracy 개선 없음
3	VGG16 & ArcFace	27 epoch 학습 기준 Training Accuracy: 100%, Test Accuracy: 77.2%

각 버전의 코드를 분석하고, **코드의 구조 및 성능 문제를 파악하는 과정**을 통해 트랙원들은 **모델 개선 방법을 탐구**했습니다. 또한, 코드 리뷰 과정에서는 **코드 일부를 빈칸으로 두고 트랙원들이 직접 채우며 복습하는 방식**으로 진행했습니다.

### 4) 트랙 커리큘럼

#### 트랙 커리큘럼

Aa 주차	≡ 학습 내용	≡ 진행 방식	≡ 교재 및 자료	≡ 실습 과제
1주차	Lecture : Basics of Convolutional Neural Networks & Deep Convolutional Models: A Case Study	대면		Version 1
2주차	Lecture : Object Detection & Special applications: facial recognition and neural style transfer	대면		Version 2
3주차	Lecture : recurrent neural network & Natural Language Processing and Word Embedding	대면		Version 3
4주차	Lecture : Sequence model and attention mechanism & Transformer Network	대면		Version 3
중간고사 휴동				
5주차	Paper review & Code review : Attention is All You Need	대면	Paper Link : <a href="https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf">https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf</a>	
6주차	Paper review & Code review : An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale	대면	Paper Link : <a href="https://arxiv.org/pdf/2010.11929">https://arxiv.org/pdf/2010.11929</a>	
9주차	프로젝트 진행	대면		
10주차	프로젝트 진행	대면		
11주차	프로젝트 진행	대면		
12주차	프로젝트 진행	대면		




소감 :

- 장\*\* : 이미지 데이터를 다루는 법과 관련 모델들에 대해서 자세히 배울 수 있었습니다. 또한 학습했던 모델들을 실제 프로젝트에도 적용해 볼 수 있어 좋았습니다.



활동사진  
(3매 이상)



<p>동아리활동과 관련하여 상기 결과보고서를 확인합니다.</p> <p>※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.</p> <p>대표학생 : 정유진 </p>	

# 동아리 활동보고서

활 동 명	심화 프로젝트 컨퍼런스 및 수료식		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 11. 27. 19:00 ~ 21:00	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 학습한 내용을 바탕으로 실용성 있는 프로젝트를 진행하여 실무 및 연구 역량을 증진한다.</p> <p>2. KHUDA 회칙 제4장 제15조 [컨퍼런스] 컨퍼런스는 회원들이 주도적으로 진행한 연구 결과를 공유하고 피드백을 받는 자리로, 팀 또는 개인 기반의 프로젝트 결과를 발표한다.</p>		
활동내용 및 소감	<p>장소 : 멀티미디어관 112호</p> <p>활동내용 :</p> <p>본 행사는 KHUDA 내에서의 심화 학습 및 프로젝트 결과 공유의 장으로, 학습 성과의 평가 및 인증을 위해 매 학기 진행해오고 있습니다. 신규 회원과 기존 회원 모두 컨퍼런스 참여를 통해 수료 조건을 충족하면 수료증을 받게 되는 체계적인 프로그램입니다.</p> <p>이번 컨퍼런스는 KHUDA의 모토인 "함께 성장하며 한계를 뛰어넘는 여정"을 마무리하는 자리입니다. 보다 더 심화된 학습을 진행한 이후에 회원들이 서로 수행한 프로젝트 결과를 공유하고, 성과를 평가하는 자리로 마련하였습니다. 수료식은 신규 회원과 기존 회원이 수료 조건에 맞춰 동아리 활동을 성실히 수행했을 경우 수료하는 방식으로 진행했습니다.</p> <ul style="list-style-type: none"> <li>- 신규 회원: ML 기초 세션, ML 토이 프로젝트, 심화 세션, 심화 프로젝트 수행을 모두 완료한 경우 수료</li> <li>- 기존 회원: 심화 프로젝트 수행을 완료한 경우 수료</li> </ul> <p>1) 컨퍼런스 진행 및 평가 방식</p> <p>이번 학기 KHUDA 회원들의 활발한 활동 덕분에 진행된 프로젝트가 매우 많았습니다. 모든 프로젝트를 컨퍼런스에서 발표하기에는 시간적 여유가 부족하여, 신규 회원과 기존 회원 사이의 발표 방식을 다르게 진행하였습니다. 이는 신규 회원에게 발표 기회를 우선적으로 제공하기 위함입니다.</p> <ul style="list-style-type: none"> <li>- 심화 세션 프로젝트: PPT를 활용한 발표로 진행하여 각 세션의 프로젝트 수행 결과를 상세히 설명하고 피드백을 받는 시간을 가졌습니다.</li> </ul>		

- **기존 회원(OB) 프로젝트:** 패널 발표 방식으로 진행되어 프로젝트 결과물을 판넬에 정리하고 자유롭게 토론하며 깊이 있는 피드백을 진행했습니다.

소프트웨어융합학과의 교수님 두 분께서 참석하셔서 컨퍼런스의 질을 높이는 심사와 피드백을 제공해 주셨습니다. 교수님들의 전문적인 지식과 경험은 참여 학생들에게 실질적인 학문적 지원을 제공해주셨으며, 추후 프로젝트 개선의 질을 한층 더 끌어올리는 데 크게 도움이 되고 있습니다.

- **참석 교수님:** 소프트웨어융합학과 조명아 교수님, 소프트웨어융합학과 박제만 교수님
- **평가 방식:** 회원 평가 점수 25 % + 교수 평가 점수 75 %
- **평가 기준은** 운영진 내부 회의를 거쳐, 프로젝트의 완성도, 창의성, 기술성, 활용성, 전달성을 균형 있게 반영할 수 있도록 설정하였습니다.

### 세부 평가 기준

- 심화 컨퍼런스 평가는 다음과 같은 평가 항목과 배점으로 구성됩니다. 우선 발표자가 해결하고자 하는 문제 상황을 명확하게 제시하는지가 20점으로 배점되어 가장 높은 비중을 차지합니다. 동일한 비중으로 문제 해결을 위한 파이프라인 설명이 20점이 배정되어 있으며, 이는 해결과정의 체계성과 논리성을 평가합니다. 학습한 내용을 실제 문제에 어떻게 적용했는지에 대한 설명과 구체적인 결과물 소개는 각각 15점씩 배정되어 있습니다. 연구의 한계점을 인지하고 향후 개선 방향을 제시하는 능력은 10점으로 평가됩니다. 발표 시 청중과의 소통을 위한 발표 전달력과 질의응답 시간에 이루어지는 Q&A 대응 능력도 각각 10점씩 배정되어 있습니다. 평가는 지도 교수진, 학생들, 그리고 운영진이 균형 잡힌 비율로 참여하여 진행되었습니다.
- 이번 컨퍼런스에서는 데이터 비즈니스 트랙, 컴퓨터 비전 트랙, 그리고 OB들이 참여한 디플로이 트랙에서 우수한 성과를 보인 팀들이 수상의 영예를 안았습니다. 이러한 평가 체계를 통해 문제 해결 능력뿐만 아니라 의사소통 능력까지 종합적으로 평가하여 각 트랙별 최고의 프로젝트를 선정하였습니다.

### 2) 주요 프로젝트 발표 주제

구분	구분 상세	발표 주제
심화 세션	금융	KHUSPI 지수 예측
	NLP	KHUrious: 경희대의 Curious한 정보를 알려줄게
		Emotify: 텍스트에 감정을 더하다
	데이터 비즈니스 & 추천시스템	미래 가계부: 내일의 돈을 오늘 설계하다
		런런이를 위한 경희대학교 주변 러닝코스 최적화
		지진 대비 추가적인 임시주거시설 추천
		NOTD(News Of The Day): 오늘의 사건록! 당일 핵심 뉴스와 사건 흐름을 한눈에 보여주는 대시보드 서비스
	CV	PhotoRealism Depth Alignment 기반

기 존 회 원		자연스러운 인생네컷 눈 감지 마! 단체 사진에서 눈 감으면 알려주는 서비스
	자율 프로젝트	CellInsight 단일 세포 분석 효율 증진 툴
	공간 분석 스터디	지하철 도심물류시스템 환경적 실효성 검증과 최적화
	Deploy Session	LINK: 저장한 링크를 키워드 정리 및 요약하는 서비스
		당하그: 당신의 하루를 그려드립니다
	핸즈온 AWS 스터디	온라인 소비자 보호를 위한 다크 패턴 감지 툴
	Apache Kafka 애플리케이션 프로그래밍 스터디	실시간 서울시 상권 데이터 분석
	자율 프로젝트 데이터사이언스 & 시각화 스터디	ATOS spotify API를 이용하는 음악 추천 시스템 최고의 항공사를 찾아라!

### 3) 교수님 피드백

두 교수님께서서는 전체적으로 동아리의 학습 시스템이 이론적 학습부터 실전 프로젝트 수행까지 단계적으로 구성된 점을 긍정적으로 평가해 주셨습니다. 이러한 일련의 과정을 통해 학생들이 자기 주도적으로 문제를 정의하고 해결해 나가는 역량을 기를 수 있다는 점을 높이 평가하셨습니다. 그리고 심화 프로젝트 발표뿐만 아니라, 자율적으로 진행된 독립 프로젝트의 존재 역시 동아리 내 연구 및 학습 문화의 성장을 보여주는 중요한 지표라고 언급하셨습니다. 교수님들께서 남겨주신 피드백을 바탕으로 한발짝 더 나아가는 동아리 KHUDA가 되도록 노력할 것입니다.

**A 교수님:** 모든 프로젝트에 참여한 학생들의 **열정과 실력**을 높이 평가하며, 특히 프로젝트들이 **실용적인 문제 해결을 목표로 한 점**을 긍정적으로 평가해 주셨습니다. 금융 데이터, 추천 시스템, 컴퓨터 비전, NLP 등의 다양한 트랙에서 진행된 연구들이 실제 산업에서 어떻게 적용될 수 있는지에 대한 질문을 던지면서, 현실적인 응용 가능성에 대한 조언을 아끼지 않으셨습니다. 또한 데이터 분석 및 머신러닝 모델을 구현하는 과정에서 **데이터 전처리, 하이퍼파라미터 튜닝, 모델 성능 개선 등의 기법을 보다 체계적으로 접근할 필요가** 있다고 조언하시며, 향후 연구 방향을 설정하는 데 도움을 주셨습니다.

**B 교수님:** 프로젝트의 **창의성과 문제 해결 능력**에 초점을 맞추어 개별 발표자들에게 세부적인 피드백을 제공해 주셨습니다. 특정 프로젝트에서 사용된 **모델 선정의 타당성, 평가 지표의 적절성, 데이터 셋 구성 방식** 등에 대해 구체적인 질문을 던지며, 연구의 신뢰도를 높이는 방법에 대해 함께 논의해 주셨습니다. 또한 AI 및 데이터 분석 프로젝트는 단순히 높은 성능을 내는 것이 중요한 것이 아니라, 실제 도메인에서 신뢰할 수 있는 결과를 도출할 수 있는지가 관건이라며 비즈니스 및 연구적 측면에서 모델의 설명력과 해석 가능성을 높이는 것이 중요하다는 점을 강조하셨습니다.



소감 :

- 최\*\* : NOTD 프로젝트에서 요약 모델을 담당했습니다. KoBART 모델을 학습시키고 타임라인에 들어갈 요약문이 보다 완성도 있을 수 있도록 성능을 높이는 데에 힘썼습니다. 마지막에는 대시보드 상단의 프론트엔드 부분도 작업하여 전체적으로 그럴듯한 프로젝트로 보이게끔 노력했습니다. 힘든 점이 있었다면.. NLP, 프론트 모두 다 처음 해보는 것이었다는 점입니다. 사실 워낙 프로젝트 경험이 없어서 어느 부분을 맡았어도 처음 해보는 것이었을 겁니다. 공부해나가는 과정이 힘들었지만 재밌었고, 팀 분위기가 너무 좋았어서 수월하게 진행할 수 있었습니다.

- 이\*\* : 기본적으로 코딩 실력이 0에 수렴할 정도였을 때 해당 동아리에 들어왔는데, 심화 세션을 진행하면서 코딩 지식이 상당히 많이 늘었다. 코드를 작성함에 있어서 가졌던 막연한 두려움이 확연히 줄어들었다. 다만 데이터 비즈니스라는 분야가 워낙 방대해, 프로젝트를 진행할 때 세션에서 배웠던 지식을 거의 쓰지 못한 부분은 많이 아쉽다. 그래도 프로젝트 진행은 너무 재밌었고, 협업할 때의 마음가짐, 프론트 지식 등 많은 것을 배웠다.

활동사진  
(3매 이상)



The poster for the KHUDA 6th DEEP DIVE CONFERENCE features a dark background with glowing star-like patterns. The title 'KHUDA 6th DEEP DIVE CONFERENCE' is prominently displayed in the center. Below the title, the date and time '2024.11.27.WED PM 7:00 ~ 9:00' and the location '멀티미디어관 112호' are listed. The poster also includes a QR code in the top right corner and several session titles in Korean, such as 'KHUSPI 지수 예측', '미래 가계부 내일의 돈을 오늘 설계하다', '오늘의 사건픽! 당일 핵심 뉴스와 사건 흐름을 한눈에 보여주는 대시보드 서비스', '지하철 도심물류시스템 환경적 실용성 검증과 최적화', 'Emotify 텍스트에 감정을 더하다', 'Depth Alignment를 통한 자연스러운 인생네트', 'ATOS Spotify API를 이용하는 음악 추천 시스템', 'LINK Link Is Not Kind', 'PeTurning 온라인 소비자 보호를 위한 다크 패턴 감지 툴', 'KHUrious 경의대의 Curious한 정보를 알려줄게!', '실시간 서울시 상권 데이터 분석', 'Cellinsight 단일 세포 분석 효율 증진 툴', and 'KHUDA 6기 활동 보러가기 >>'. A small asterisk is located at the bottom right of the poster.

03

◆  
**CellInsight**  
단일 세포 분석 효율 증진 툴  
데이터 엔지니어링 | 추천 시스템  
류여진, 이하영

◆  
**지하철 도심물류시스템**  
환경적 실효성 검증과 최적화  
공간분석 | 데이터 비즈니스  
정난주, 정소연, 정유진, 조준영

◆  
**LINK**  
Deploy Session | NLP  
박현준, 정유진, 김채영

◆  
**온라인 소비자 보호를 위한 다크 패턴 감지 툴**  
데이터 비즈니스 | 데이터 엔지니어링  
김건형, 김민아, 박상영, 유혜지, 허윤지

◆  
**실시간 서울시 상권 데이터 분석**  
데이터 엔지니어링  
김건형, 유혜지, 윤소은, 이수민, 이재호, 홍민혁, 황종훈

◆  
**ATOS**  
Spotify API를 이용하는 음악 추천 시스템  
데이터 비즈니스 | 추천 시스템  
Dania

◆  
**당하그**  
Deploy  
김수진

◆  
**최고의 항공사를 찾아라!**  
데이터 시각화  
권서연, 서지은, 이두원

**KHUDA** 5th DEEP DIVE  
**CONFERENCE**

2024.11.27 WED PM 7:00 ~ 9:00  
멀티미디어관 210호

03

◆  
**오늘의 사건록!**  
당일 핵심 뉴스와 사건 흐름을 한눈에 보여주는 대시보드 서비스  
NLP | 데이터 비즈니스 | 추천 시스템  
김정운, 장하늘, 김채영, 최예지

◆  
**지진 대비 추가적인 임시주거시설 추천**  
데이터 비즈니스 | 추천 시스템  
최진영, 송인혁, 정소연, 박정현, 채진기

◆  
**눈 감지 마!**  
단체 사진에서 눈 감으면 알려주는 서비스  
CV  
김민서, 김예람, 장준필, 이하영

◆  
**PhotoRealism**  
Depth Alignment 기반 자연스러운 인생네컷  
CV  
문강민, 신수현, 이승준, 이찬, 이하영

◆  
**Emotify : 텍스트에 감정을 더하다**  
NLP  
김동혁, 김혜란, 이동진, 장영주

◆  
**KHUrrious**  
경희대의 Curious한 정보를 알려줄게!  
NLP | 데이터 엔지니어링  
권도현, 박찬희, 오윤진

◆  
**런런이를 위한 경희대학교 주변 러닝코스 최적화**  
데이터 비즈니스  
박지연, 윤예준, 전지현, 정유진

◆  
**KHUSPI 13 ETF선정 및 예측**  
일별 뉴스 데이터 감성 점수 기반  
NLP | 금융  
선현석, 신경수, 안형민, 임주현, 장서연, 정혜인, 김형준

◆  
**미래 가게부: 내일의 돈을 오늘 설계하다**  
데이터 비즈니스  
이재희, 이준영, 김채영, 허채은

KHUDA 경희대에서  
자세히 보기

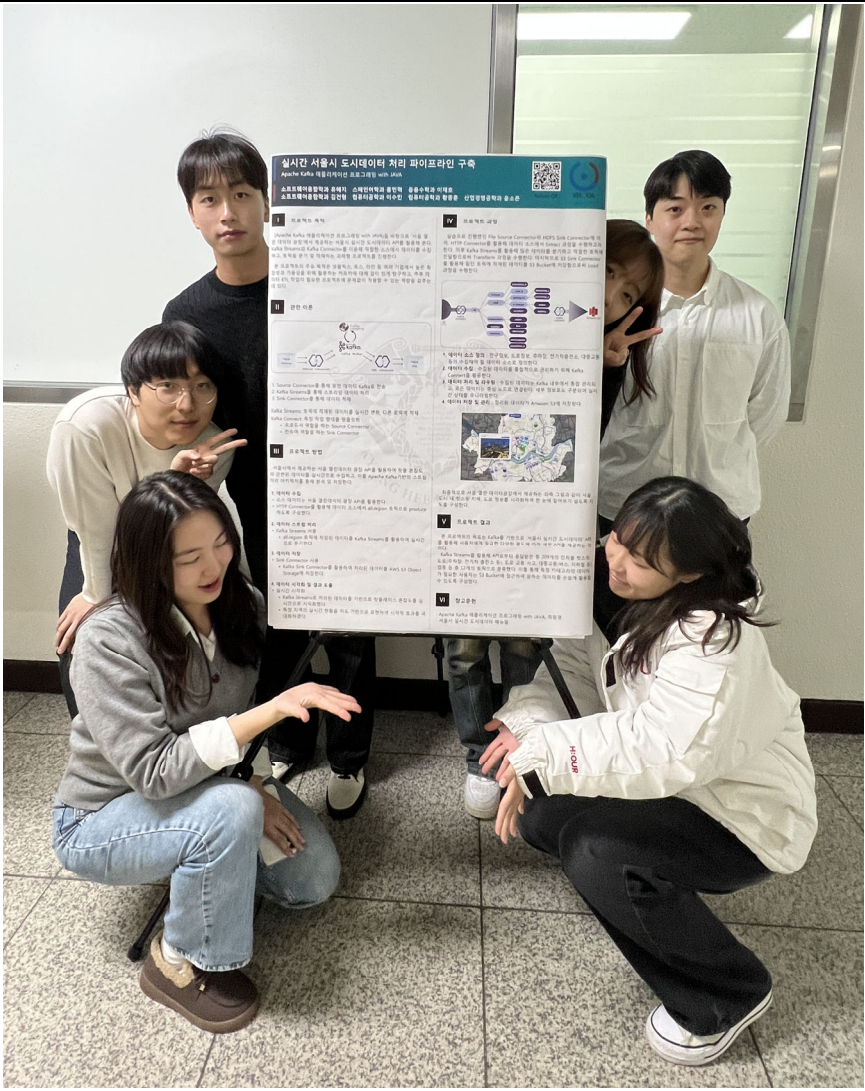


6기 활동 노선  
바로가기

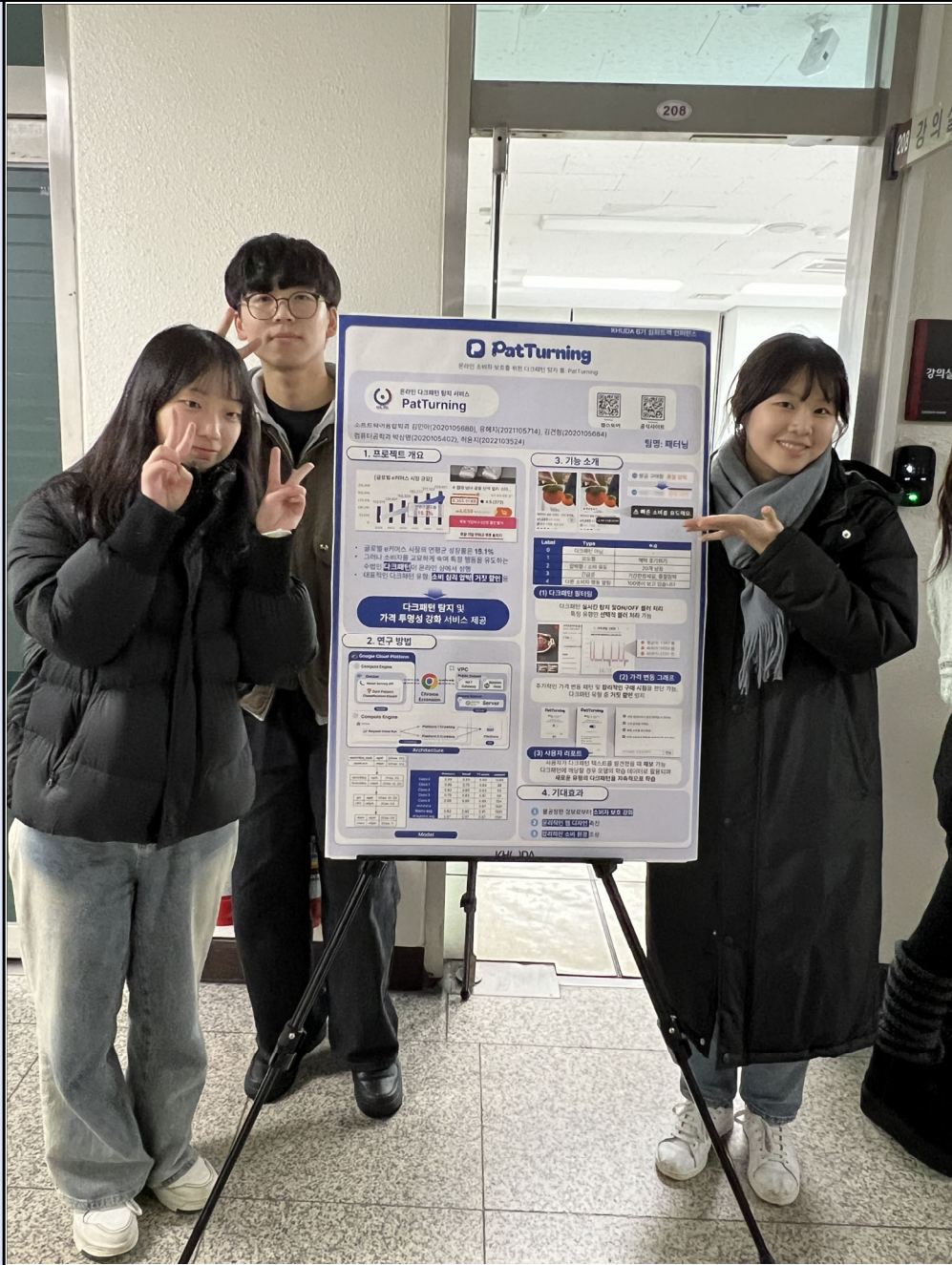


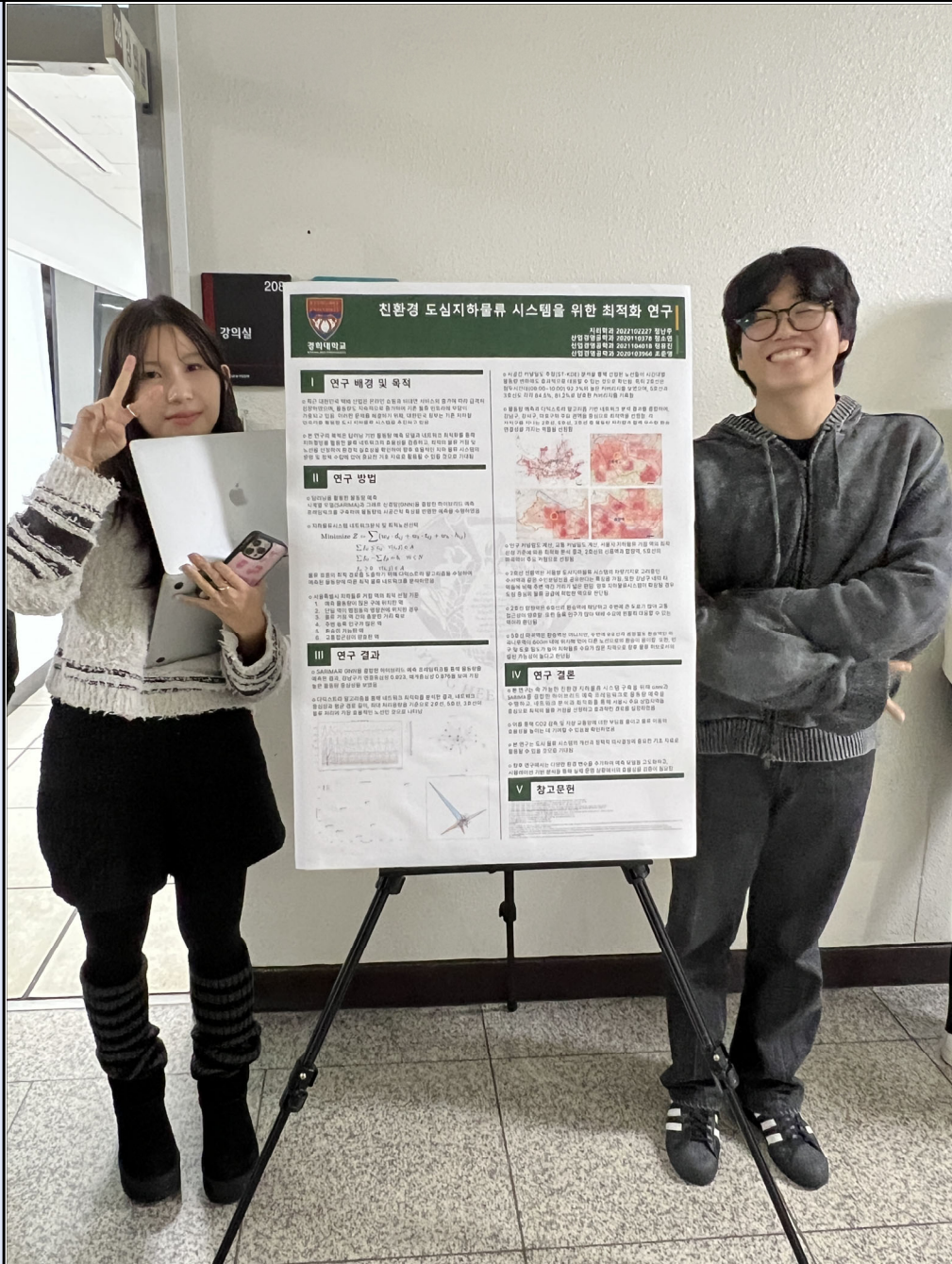















### 친환경 도심지하물류 시스템을 위한 최적화 연구

김민준 교수 2022102227 정남우  
신민영 학생부 2020110178 조승우  
김민준 학생부 2022104016 정남우  
김민준 학생부 2022013066 조승우

#### I 연구 배경 및 목적

우리는 친환경적이며 효율적인 운송 수단인 지하물류 시스템을 통해 도시의 교통 혼잡을 완화하고, 환경 오염을 줄이고, 시민의 편의성을 높이기 위해 연구를 진행하였다. 본 연구의 목적은 친환경적이며 효율적인 지하물류 시스템을 설계하고, 이를 최적화하여 도시의 교통 혼잡을 완화하고, 환경 오염을 줄이고, 시민의 편의성을 높이기 위한 것이다.

#### II 연구 방법

이 연구를 위한 방법론은 다음과 같다. 첫째, 지하물류 시스템의 구조를 설계하고, 둘째, 지하물류 시스템의 운영을 최적화하기 위한 알고리즘을 개발하고, 셋째, 개발된 알고리즘을 시뮬레이션하여 그 성능을 평가한다.

Minimize  $Z = \sum_{i,j} (w_{ij} \cdot d_{ij} + c_{ij} \cdot w_{ij} \cdot h_{ij})$

Subject to  $\sum_{j \in N} w_{ij} = 1, \forall i \in N$

$\sum_{i \in N} w_{ij} = 1, \forall j \in N$

여기서  $w_{ij}$ 는 노드  $i$ 에서 노드  $j$ 로 가는 물류량의 비율,  $d_{ij}$ 는 노드  $i$ 와 노드  $j$  사이의 거리,  $c_{ij}$ 는 노드  $i$ 에서 노드  $j$ 로 가는 단위 물류량당 비용,  $h_{ij}$ 는 노드  $i$ 에서 노드  $j$ 로 가는 단위 물류량당 환경 오염량이다.

#### III 연구 결과

본 연구의 결과는 다음과 같다. 첫째, 지하물류 시스템의 구조를 설계하고, 둘째, 지하물류 시스템의 운영을 최적화하기 위한 알고리즘을 개발하고, 셋째, 개발된 알고리즘을 시뮬레이션하여 그 성능을 평가한다. 결과적으로, 본 연구의 결과는 도시의 교통 혼잡을 완화하고, 환경 오염을 줄이고, 시민의 편의성을 높이는 데 크게 공헌하였다.

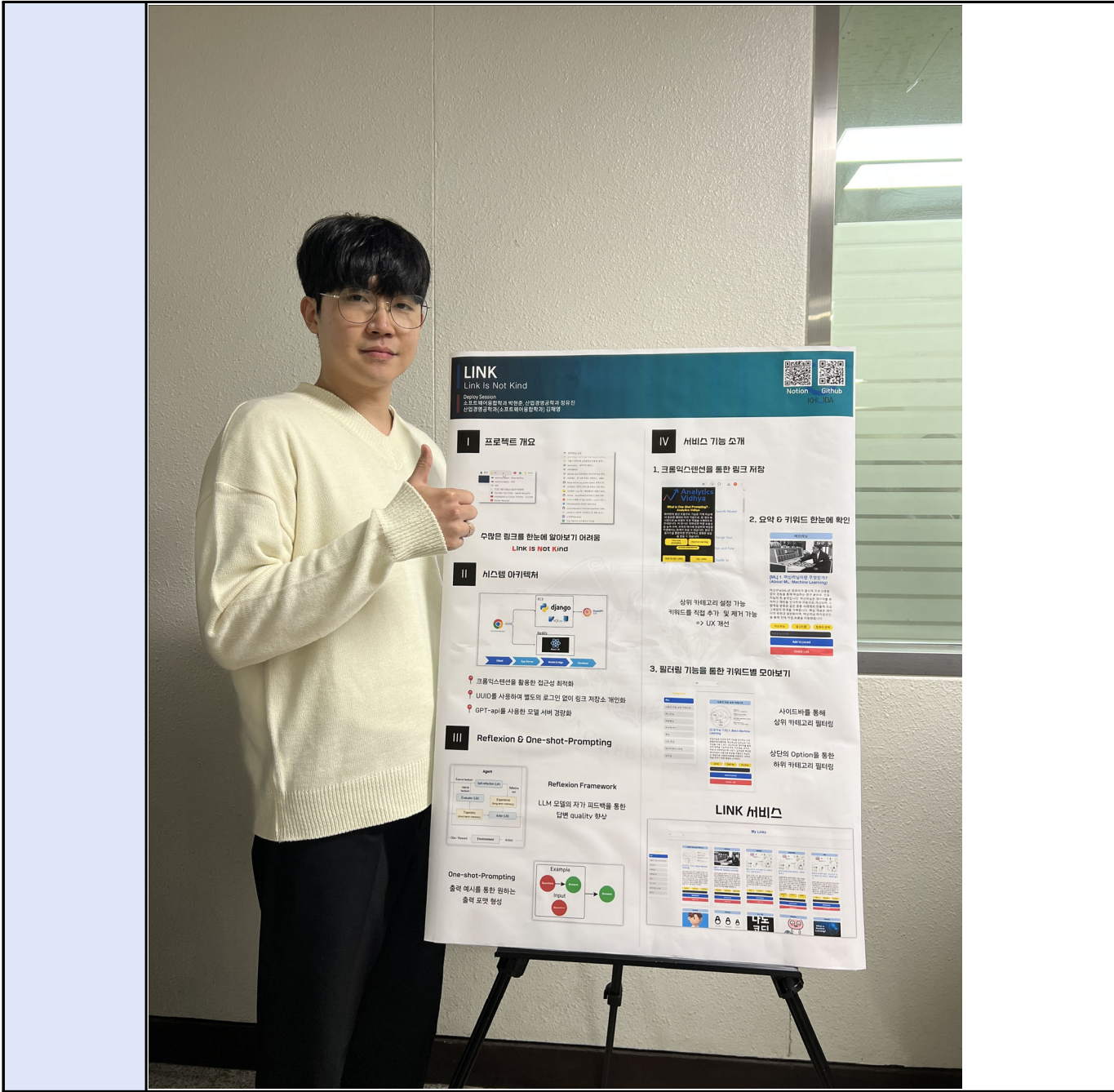
#### IV 연구 결론

본 연구의 결론은 다음과 같다. 첫째, 지하물류 시스템의 구조를 설계하고, 둘째, 지하물류 시스템의 운영을 최적화하기 위한 알고리즘을 개발하고, 셋째, 개발된 알고리즘을 시뮬레이션하여 그 성능을 평가한다. 결과적으로, 본 연구의 결과는 도시의 교통 혼잡을 완화하고, 환경 오염을 줄이고, 시민의 편의성을 높이는 데 크게 공헌하였다.

#### V 참고문헌

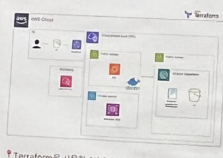
본 연구에 참고한 문헌은 다음과 같다. 김민준, 정남우, 조승우, 2022. 친환경적이며 효율적인 지하물류 시스템 설계 및 최적화 연구. 서울대학교 학위논문. 정남우, 김민준, 조승우, 2021. 친환경적이며 효율적인 지하물류 시스템 설계 및 최적화 연구. 서울대학교 학위논문. 조승우, 김민준, 정남우, 2020. 친환경적이며 효율적인 지하물류 시스템 설계 및 최적화 연구. 서울대학교 학위논문.

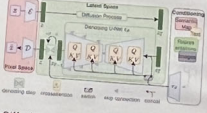




**당하고(당신의 하루를 그려드립니다)**  
 Stable diffusion을 이용한 회고 서비스  
 Deploy Session 컴퓨터공학과 김수진(2022105673)

**I 프로젝트 개요**  
 글쓰기, 회고 등 회고 스티디(모임)에 대한 수요 증가  
 But 현재 회고를 위한 서비스는 많이 없어 주로 노션, 슬랙 등을 이용

**II 시스템 아키텍처**  
  
 Terraform을 사용한 AWS 리소스 자동화  
 EC2를 통해 클라이언트 요청 처리  
 SageMaker를 이용해 AI 모델 학습 및 추론 수행

**III Stable-Diffusion**  
  
**Stable Diffusion**  
 텍스트를 입력받아 이미지를 생성하는 딥러닝 모델로, Latent Diffusion Model(LDM)을 기반으로 텍스트 표현력(문맥, 구체성)의 이미지를 생성  
**특징**  
 이미지(문맥)를 latent space(잠재 공간)에서 처리해 GPU 메모리 사용량을 줄이고, 텍스트 표현력에 고품질 이미지를 생성

**IV 기능 소개**  
 1. 하루 목표 설정  
 오전 10시 이전까지 하루다짐 작성  
 2. 회고 작성  
 매일 작성 이전까지 회고 작성  
 3. 이미지 생성  
 Stable-diffusion을 이용해 작성한 회고를 기반으로 이미지 생성  
**Stable Diffusion 프롬프트 엔지니어링**  
 페르소나와 생성할 이미지 스타일을 지정하고 디테일을 강화하여 서비스에 적합한 이미지를 생성하도록 함





**CellInsight**  
단일 세포 분석 (Single Cell Analysis) 효율 증진  
유전체생물학과 생명인  
소분배아생물학과 이라영

**I. 프로젝트 목적**  
Single Cell RNA-Seq/ATAC-seq, 단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영  
단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영

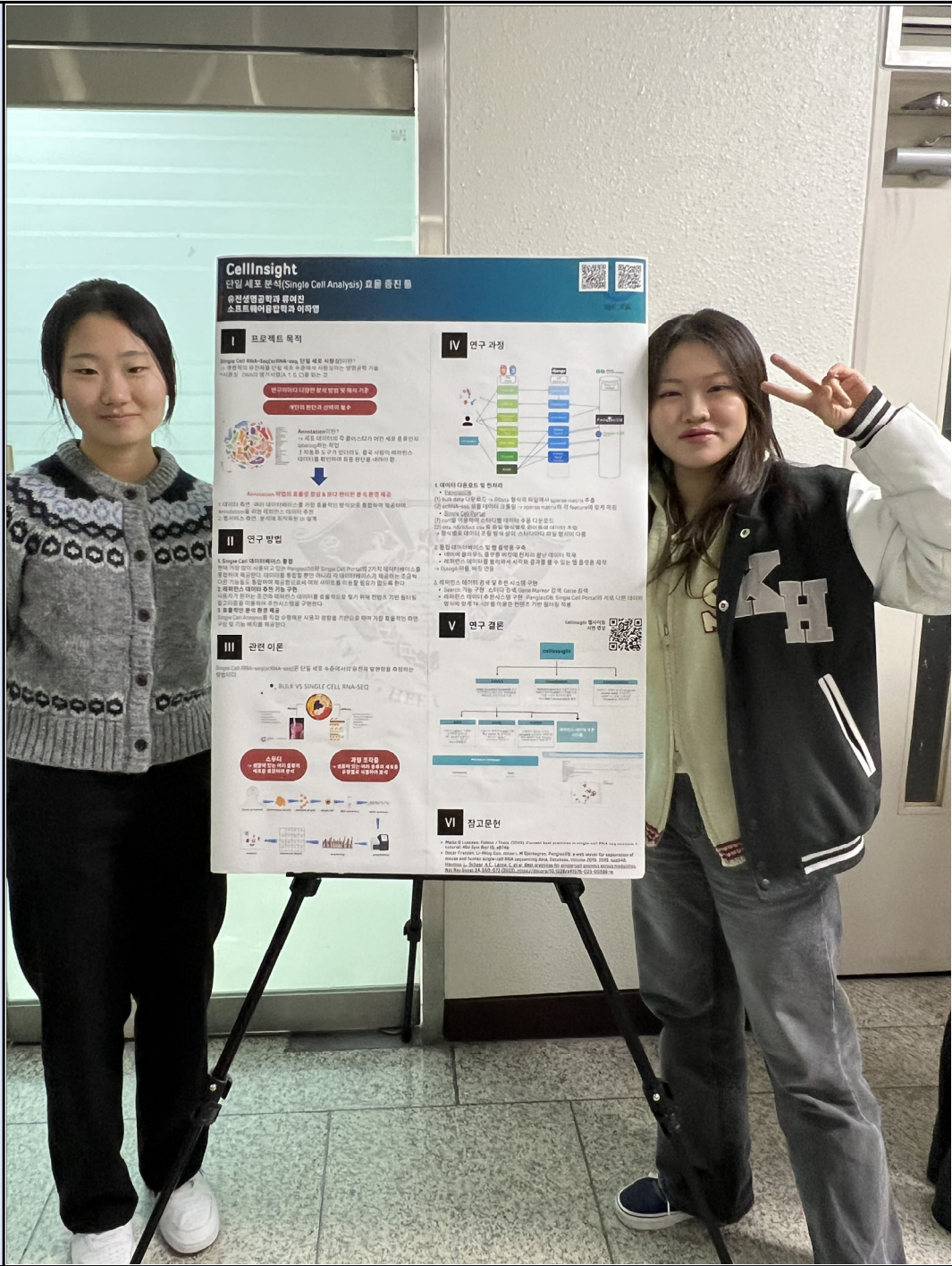
**II. 연구 방법**  
1. Single Cell RNA-Seq/ATAC-seq, 단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영  
단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영

**III. 관련 이론**  
Single Cell RNA-Seq/ATAC-seq, 단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영  
단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영

**IV. 연구 과정**  
1. 데이터 수집 및 전처리  
2. 데이터 분석 및 시각화  
3. 데이터 해석 및 결론 도출

**V. 연구 결론**  
1. 연구 목적 달성  
2. 연구 결과 요약  
3. 연구의 의의와 한계

**VI. 참고문헌**  
1. Single Cell RNA-Seq/ATAC-seq, 단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영  
단일 세포 분석 방법론의 발전  
유전체생물학과 생명인, 소분배아생물학과 이라영







동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

# 동아리 활동보고서

활 동 명	스터디 - 공간 분석 스터디		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 8. 18. ~ 2024. 9. 15. 매주 1회 22:00 ~ 23:30	비고	경희대학교 동아리 학술제 CAAS 학술부문 대상 수상
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다.</p> <p>2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>3. KHUDA 회칙 제4장 16조 [스터디] 목적: 스터디는 특정 주제에 대해 관심 있는 회원들이 자율적으로 모여, 해당 주제에 대해 학습 후 관련 경험을 쌓는다.</p>		
활동내용 및 소감	<p>장소 : 비대면 진행</p> <p>활동내용 :</p> <p>공간 분석과 관련된 최신 연구 동향을 파악하고, GIS, R, Python 등의 도구를 활용한 데이터 분석 및 논문 작성을 위한 심화 학습을 진행하기 위한 스터디입니다.</p> <p>공간 분석에 대한 깊이 있는 이해를 바탕으로, 실습을 통해 이론을 적용하며 연구 역량을 강화하고자 본 스터디를 개설하였습니다. 특히 GIS, R, Python을 활용한 딥러닝 및 최적화 기법을 연구하고, 이를 바탕으로 학술 논문 작성을 진행하는 것을 목표로 삼았습니다.</p> <p><b>1) 스터디 진행 방식</b></p> <p>스터디 커리큘럼은 스터디원들의 의견을 최대한 반영하여 유연하게 구성하고 주차별 계획을 성실히 수행했습니다.</p> <p>스터디 활동 외에 연구와 논문 작성을 원하는 스터디원들은 추가적인 스터디 시간을 가졌습니다. 기본적으로 GIS 툴과 Python+R 학습을 두 개의 큰 축으로 설정하여 운영했습니다.</p> <p>GIS 툴 학습의 경우, GIS 학회에서 실습 자료를 체계적으로 제공하고 있어서, 이를 최대한 활용하여 학습했습니다. 초반에는 GIS 툴을 집중적으로 실습하는 시간을 가졌고, 후반부에는 Python과 R을 활용한 공간 분석 및 딥러닝을 중심으로 학습했습니다.</p> <p>스터디 참여 대상은 서비스 개발보다는 데이터 분석에 관심이 있는 사람, 실습 진행에 적극적으로 참여하고 문제 해결 및 토론을 즐기는 사람, GIS 툴과 Python/R을 활용한 분석을 배우고 싶은 사람이라면 누구나 참여할 수 있도록 했습니다.</p> <p><b>2) 주차별 계획</b></p>		



Aa 주차	:≡ 주차주요 개념	≡ 학습목표
<b>1</b> 1주차	GIS	GIS 툴 기본 튜토리얼 및 사용법 학습 GIS 기본 개념 및 데이터 처리 실습 (Shapefile, GeoJSON 등)
<b>2</b> 2주차	GIS	GIS 데이터 처리 실습 (join) 결과 출력 실습
<b>3</b> 3주차	GIS	벡터 / 래스터 데이터 실습
<b>4</b> 4주차	R Python	Python을 사용한 GIS 데이터 처리 및 시각화 (Geopandas, Matplotlib 등)
<b>5</b> 5주차	R Python	R을 사용한 GIS 데이터 분석 (sf 패키지 등)
<b>6</b> 6주차	R Python	Python을 사용한 딥러닝 모델 구축 및 실습 (간단한 예제 프로젝트)
<b>7</b> 7주차	github R Python	GitHub 올리기

### 3) 주차별 보고

본 스터디는 **매주 발제 담당자를 지정하여** 진행하였습니다. 발제자는 **해당 주차의 학습 내용을 사전에 충분히 연구한 후, 스터디원들이 실습을 원활하게 따라올 수 있도록 설명하는 역할**을 맡았습니다. 실습 중 발생할 수 있는 질문에 대비하여 미리 답변을 준비하였으며, 매주 세션이 끝나기 전에는 질의응답(Q&A)과 토론 시간을 마련하여 학습 내용을 더욱 심화할 수 있는 시간을 가졌습니다.

특정 주제에 대해 보다 깊이 있는 탐구가 필요할 경우에는 스터디원들이 돌아가며 심화 발제를 준비하여 공유하는 방식으로 학습했습니다. 이 과정을 통해 단순한 실습을 넘어서, 공간 분석에 대한 이해를 체계적으로 확장할 수 있었습니다.

스터디의 운영 방식은 자율적이면서도 체계적으로 구성하여, 스터디원들이 능동적으로 학습에 기여하고 적극적으로 토론에 참여할 수 있도록 노력했습니다. 본 스터디를 통해 공간 분석 역량을 효과적으로 향상시키고, 실무 적용 가능성을 높이는 데 기여했습니다.

차수	담당자	보고 링크	참고자료 링크
1차시(OT)	정난주	공간분석스터디 1주차: QGIS 다운로드 - <a href="https://goodnanju.tistory.com/m/14">https://goodnanju.tistory.com/m/14</a>	<a href="https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1515">https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1515</a>
2차시	조준영	공간분석스터디 2주차:  [QGIS] 2주차	<a href="https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1519">https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1519</a> <프로젝트/맵 탐색/속성 툴바 (1)~(3)>
3차시	정유진	2주차 (notion.site)	<a href="https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1522">https://compas.lh.or.kr/studyinfo?brdTypeCd=23&amp;pageIndex=7&amp;pageSize=10&amp;searchKey=both&amp;searchText=&amp;totalCount=118&amp;brdArtclNo=1522</a> <속성편집 ~ 지오파키지 구축_2주차 2번째 XYZ파일 구축은 거의 안쓰는 거라 빼도 될 것 같아요! (1,3,4,5만 !!)>
4차시	배아람	<a href="https://baram1ng.notion.site/4-afeec982c94b452facdf717b050b3570?pvs=4">https://baram1ng.notion.site/4-afeec982c94b452facdf717b050b3570?pvs=4</a>	 COMPAS COMPAS <3주차1,2,3>
5차시	정소연	<a href="https://cyber-gasoline-abf.notion.site/5-102777792458800fa756c5f593b83acf?pvs=4">https://cyber-gasoline-abf.notion.site/5-102777792458800fa756c5f593b83acf?pvs=4</a>	 urbanites [QGIS 실습 06] 인접성 분석 1 - 버퍼 (buffer)

#### 4) 주요 성과 및 결과물

- **프로젝트:** '친환경 도심 지하 물류 시스템을 위한 최적화 연구' 프로젝트를 수행했습니다.
- **성과:** **경희대학교 동아리 학술제 CAAS 학술 부문 대상**을 수상했습니다.

#### - 프로젝트 개요:

최근 대한민국의 택배 산업 성장에 따라 물류 처리 능력의 한계가 점차 드러나고 있습니다. 이에 따라, 서울시 내 지하철을 활용한 도심 물류 시스템이 지속 가능성과 효율성을 고려한 해결책으로 주목받고 있습니다.

본 연구에서는 딥러닝을 활용한 물동량 예측 및 네트워크 분석을 통해 지하 물류 시스템의 실효성을 검증하고, 이를 바탕으로 탄소 배출 절감 효과를 분석하였습니다. 또한 현재 지정된 3호선 외에 향후 최적의 물류 거점으로 활용될 가능성이 높은 역을 탐색하는 연구를 수행했습니다.

#### - 연구 측면에서의 기여:

1. 물류 네트워크 최적화 및 지하 물류 시스템 효율성 분석
2. 탄소 절감 효과 분석 및 친환경 물류 시스템 구축 방향 제시
3. 서울시 물류 정책 및 운영 전략 수립을 위한 기초 자료 제공

#### 소감 :

**정\*\*:** 처음엔 단순히 공간 분석을 배우고 싶어서 시작했거든요. 근데 스터디에서 배운 걸 실제로 적용해서 학술제에서 대상까지 받게 될 줄은 몰랐어요 ㅎㅎ 스터디에서 했던 실습 덕분에 데이터를 다루는 방식이 익숙해졌고, 공간 데이터를 분석하는 과정에서 중요한 요소들을 빠르게 캐치할 수 있었어요!! **개인적으로 QGIS 지도 시각화를 통해서 인사이트를**

도출하는 방법을 배운 게 정말 큰 도움이 됐어요. 그냥 개념만 익힌 게 아니라 실제로 써 먹을 수 있는 기술을 배웠다는 게 가장 뿌듯합니다!!^^

**정\*\*:** 원래 공간 데이터를 다루는 직무에 관심이 있어서 참여했는데, 실제로 써보니까 이게 생각보다 활용도가 엄청 크다는 걸 느꼈어요. 그냥 지도 데이터라고 해서 어렵게만 생각했는데 분석 방법을 배우니까 '이걸 회사에서도 바로 써먹을 수 있겠다' 싶더라고요?! QGIS랑 파이썬을 같이 써서 공간 데이터를 가공하는 법을 익힌 것도 좋았고, 최고의 스터디원들과 함께 한 덕분에 앞으로 공간 분석을 더 깊이 공부해볼 동기부여가 됐어요ㅎ!

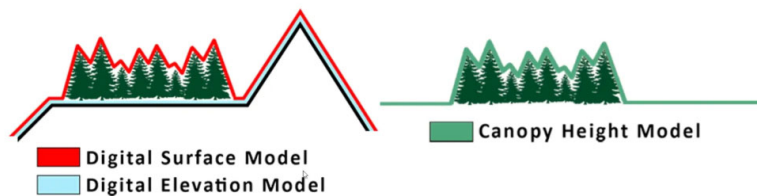
**배\*\*:** 스터디원들 앞에서 발표하려고 자료를 정리하는 과정을 거치면서, 그냥 따라 하기만 했을 때보다 개념이 훨씬 더 잘 이해됐어요. 특히 공간 분석 툴을 하나하나 직접 만져보면서, '아, 이렇게 조합하면 특정 지역의 특성을 분석할 수 있구나' 하는 감이 왔어요. 심화 발표를 준비하면서 다른 스터디원들도 같이 흥미를 느껴주니까 엄청 보람차더라고요? 혼자 공부했으면 절대로 이렇게 깊이 있게 배우지는 못했을 것 같아요.

**조\*\*:** 데이터 분석을 하긴 했었지만, 공간 데이터를 다룬 건 처음이었다. 평소에 그냥 엑셀이나 테이블 데이터만 분석했었는데 위치 정보가 들어가니까 분석 방식이 완전히 달라지는 점이 인상 깊었다. 공간적 패턴을 찾는 과정이 재밌었다. 처음에는 '이게 무슨 의미가 있을 까?' 했던 데이터들이, 분석을 해보니 지역별 특징이 확실히 보이는 게 무척이나 신기했다. 이걸 실제 정책이나 마케팅에 적용하면 정말 강력한 도구가 될 것 같다.

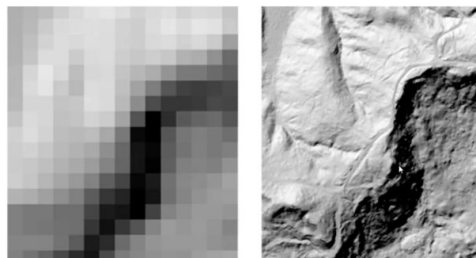
**정\*\*:** 처음에는 공간 분석이라는 개념이 생소해서 걱정했는데 스터디를 하면서 점점 흥미가 생겼어요. 특히 지도 데이터를 다루면서 실제로 시각적으로 변화를 확인하는 게 너무 신기했어요. 그냥 숫자로만 보던 데이터를 지도 위에서 색깔이나 패턴으로 나타내니까 훨씬 직관적으로 이해할 수 있더라고요. 공간 데이터를 활용하면 할 수 있는 게 정말 많다는 걸 알게 됐어요.

## DEM

- 수치표고모델로, 수치지면자료를 이용하여 격자형태(레스터 데이터)로 제작한 지표모형 구조물들의 높이는 제외한 형태의 모형
- Digital surface
  - 구조물 높이에 대한 기준이 다르게 둘 수 있는데, 이는 아래와 같다.



- 음영기복도
  - 태양의 위치를 내가 원하는 위치에 놓고 그에 따른 음영을 확인하는 것



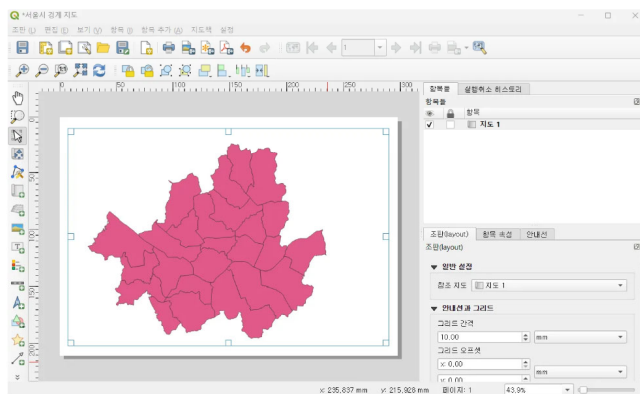
A 30m resolution DEM on the left and a high resolution 1m lidar derived DEM hillshade on the right

## 프로젝트/맵 탐색/속성 툴바 (1)

- 프로젝트 툴바



- 모든 데이터는 프로젝트 단위로 되어있음.
- 새 프로젝트, 프로젝트 열기, 저장, 인쇄 조판(프로젝트에서 작성한 데이터를 잘 디자인해서 인쇄용으로 제작하는), 스타일관리자(지도 디자인과 관련된 요소)
- 새 인쇄 조판>맵추가



- 새 인쇄 조판>아틀라스
  - 여러개의 지도가 있을 때 하나씩 볼수있게 해줌
- 스타일관리자







그림9. 교통 Kernal Density Estimation

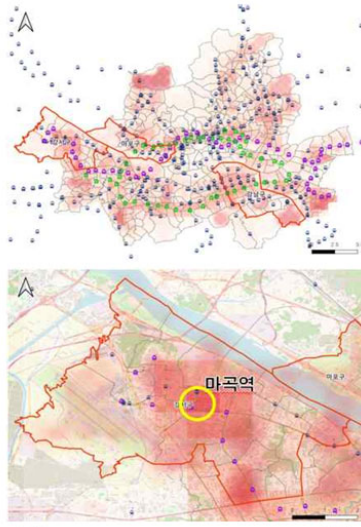
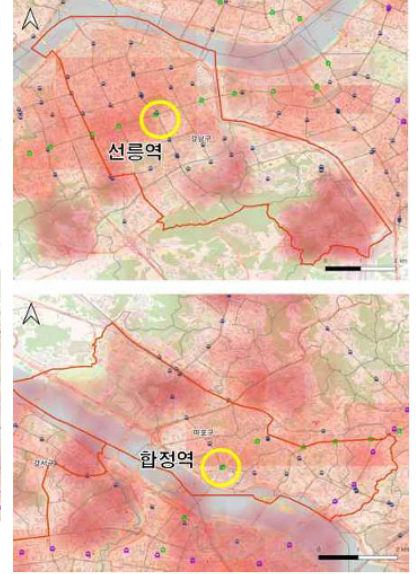


그림10. Kernal Density Estimation 중점  
및 최적역 선정





동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

## 동아리 활동보고서

활 동 명	스터디 - 핸즈온 AWS 스터디		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 7. 31. ~ 2024. 9. 2. 매주 수요일 13:00 ~ 15:00	비고	경희대학교 동아리 학술제 CAAS 학술부문 최우수상 수상
활동목적	1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다. 2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다. 3. KHUDA 회칙 제4장 16조 [스터디] 목적: 스터디는 특정 주제에 대해 관심 있는 회원들이 자율적으로 모여, 해당 주제에 대해 학습 후 관련 경험을 쌓는다.		
활동내용 및 소감	<p>장소 : 경희대학교 국제캠퍼스 정문 건너편 카페 CAFE GATE</p> <p>활동내용 :</p> <p>본 스터디는 AWS(Amazon Web Services)의 주요 기능을 실습 위주로 학습하며, 클라우드 환경에서의 실무 역량을 키우는 것을 목표로 운영하였습니다. 스터디원들은 매주 학습 범위를 나누어 발제자로서 내용을 발표하고, 실습을 통해 직접 적용해보는 방식으로 진행했습니다.</p> <p>1) 운영 방식</p> <p>1. 주차별 발제 및 실습</p> <p>매주 할당된 학습 내용을 적절히 분배하여 스터디원들이 각자 맡은 부분을 발제했습니다. 이를 통해 AWS의 다양한 서비스와 개념을 효율적으로 학습할 수 있도록 했습니다.</p> <p>2. 심화 학습 및 실습 주도 발표</p> <p>주차별로 번갈아가며 한 명의 스터디원이 교재 외 추가적인 심화 자료를 조사하여 발표하고, 실습을 준비하였습니다. 이를 통해 기본 개념을 넘어서 실무 활용도를 높이는 학습이 이루어졌습니다.</p> <p>3. 대면 진행 &amp; 협업 강화</p> <p>스터디장은 보다 원활한 토론과 실습 지원을 위해 대면 진행을 결정하였으며, 이를 통해 실시간 질의응답 및 피드백이 가능하도록 운영했습니다.</p> <p>2) 주차별 계획 및 스터디 기록</p>		

Aa 주차	☰ 주요 개념	☰ 학습 내용	👤 심화자료 및 실습 준비자
<b>1</b> 1주차	클라우드 컴퓨팅 Amazon EC2 네트워킹 Amazon VPC	1장. AWS란 1.1 클라우드 컴퓨팅(p18~) 1.2 AWS 서비스 1.3 [실습] AWS 가입하기 - 프리 티어 2장. AWS 컴퓨팅 서비스 2.1 AWS 컴퓨팅 서비스 2.2 Amazon EC2 소개 2.3 [실습] AMazon EC2 인스턴스 배포 및 접근하기(~p70) 3장. AWS 네트워킹 서비스 3.1 네트워킹이란 3.2 AWS 네트워킹 소개 3.3 Amazon VPC 소개 3.4 [실습] Amazon VPC로 퍼블릭 및 프라이빗 서브넷 구성하기(~p128)	유혜지 김건형 허윤지 박상영
<b>2</b> 2주차	부하분산 Amazon ELB CloudFormation 스토리지 Amazon EBS Amazon S3	4장. AWS 부하분산 서비스 4.1 Amazon ELB 기능 소개 4.2 [실습] ALB과 NLB를 이용한 로드 밸런싱 구성하기(~p192) 5장. AWS 스토리지 서비스 5.1 스토리지 개요 5.2 스토리지 서비스 및 주요 기능 5.3 Amazon EBS 5.4 Amazon S3 5.5 [실습] 다양한 AWS 스토리지 서비스 구성하기 (~p237)	유혜지 김건형 허윤지 박상영
<b>3</b> 3주차	데이터베이스 RDS Aurora DynamoDB ElasticCache 고급 네트워킹 DNS Amazon Route 53 CDN CloudFront	6장. AWS 데이터베이스 서비스 6.1 데이터베이스와 DBMS 6.2 AWS 데이터베이스 서비스 6.3 [실습] 웹 서버와 Amazon RDS 연동하기 (~p284) 7장. AWS 고급 네트워킹 서비스 7.1 DNS란 7.2 Amazon Route 53 서비스 7.3 CDN이란 7.4 Amazon CloudFront란 7.5 [실습] Amazon CloudFront로 CDN 서비스 구성하기 (~p335)	유혜지 김건형 박상영
<b>4</b> 4주차	IAM AutoScaling	8장. AWS IAM 서비스 8.1 배경 소개 8.2 AWS IAM 8.3 [실습] AWS IAM 사용자 생성 및 정책, 역할 동작 확인하기 9장. AWS 오토 스케일링 서비스 9.1 스케일링 9.2 AWS 오토 스케일링 서비스 9.3 [실습] Amazon EC2 오토 스케일링 구성하기 (~p413)	유혜지 허윤지 박상영
<b>5</b> 5주차	실습	10장. 워드프레스 10.1 워드프레스 소개 10.2 [실습] 워드프레스 구성하기 11장. 워드프레스 이중화 11.1 실습 소개 11.2 [실습1] AWS 서비스를 활용한 워드프레스 구성하기 11.3 [실습2] 확장성과 안정성을 고려한 워드프레스 구성하기 (~p475)	
<b>6</b> 6주차	실습	structuring architecture	
<b>7</b> 7주차	실습	building aws service one by one	
<b>8</b> 8주차	실습	implementing architecture using IaC (terraform, cloudformation, etc)	



### 3) 주요 성과 및 결과물

- **프로젝트:** [PatTurning] - 온라인 다크패턴 탐지 서비스를 제작했습니다.

- **성과:** 경희대학교 동아리 학술제 CAAS 학술 부문 최우수상을 수상했습니다.

#### - 프로젝트 개요:

서비스명 PatTurning은 "다크패턴을 감지하고 필터링하여 소비자를 보호한다"는 의미를 담고 있습니다. 본 프로젝트는 오픈소스 기반으로 개발되었으며, 기술의 투명성과 확장성을 고려하여 설계되었습니다. 실무 환경에서 클라우드 기술(AWS, GCP 등)을 적용한 프로젝트 경험 확보를 목표로 했습니다.

#### - 기술 스택 및 구현 내용:

1. 다크패턴 웹 크롤링 및 데이터 수집
2. 다크패턴 데이터 라벨링 및 AI 모델 학습
3. 모델 서빙 및 API 구축
4. 서버 및 데이터베이스(AWS 기반)
5. 크롬 익스텐션 개발 및 배포
6. GCP 환경을 활용한 PatTurning 서비스 배포

아래 QR 코드에 접속하여 다크 패턴 감지 서비스를 이용하실 수 있습니다.



PatTurning

웹스토어



PatTurning

공식사이트

#### 소감 :

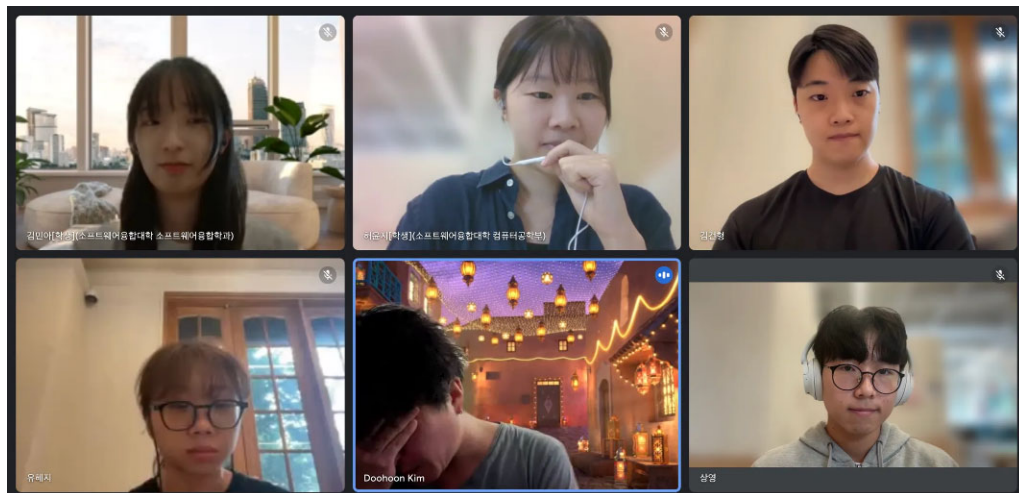
**김\*\*:** AWS가 워낙 방대해서 혼자 공부하기 어려웠는데, 스터디 덕분에 체계적으로 배울 수 있었어요. 처음엔 콘솔에서 클릭 몇 번으로 서버를 띄우는 게 신기했는데 점점 IAM 설정, 서버리스, 배포 자동화 같은 것도 배우면서 '아 이래서 클라우드가 대세구나' 싶었어요. 저는 프로젝트하면서 진짜 AWS 환경에서 배포까지 해본 게 가장 큰 경험이었습니다!! 많은 것을 배우고 얻어갑니다~~

**박\*\*:** AWS를 조금 알고 있긴 했지만, 저는 EC2나 S3 같은 기본적인 것들만 써봤거든요. 이번 스터디에서는 네트워크 구성부터 보안 설정, CI/CD까지 실제로 많이 쓰이는 기능들을 다뤄볼 수 있어서 진짜 유익했어요. 무엇보다 프로젝트를 하면서 이론이 실무나 실제 프로젝트에서 어떻게 적용되는지 확실히 감이 잡혔어요. 이제는 혼자서도 웬만한 서비스는 올릴 수 있을 것 같아요!

**허\*\*:** 책이나 강의만 보고 공부할 땐 그냥 그런가 보다 했던 개념들이 직접 AWS에서 따라 해

보니까 완전히 다르게 느껴졌어요. 특히 IAM 권한 설정 같은 부분에서 '아, 이게 왜 중요한지' 몸소 깨닫게 됐습니다..ㅎㅎ 혼자 했으면 아마 헤매다가 포기했을 수도 있는데, 스터디원들이랑 같이 하니까 문제 생길 때 바로바로 해결할 수 있어서 끝까지 따라갈 수 있었어요~!

유\*\*: 이론만 배우는 게 아니라 직접 실습을 병행해서 좋았어요~ 프로젝트를 하면서 AWS의 다양한 서비스를 조합하는 방법을 배울 수 있었던 점도 좋았어요. GCP랑 AWS를 비교하면서 사용하는 것도 재밌었어요. 실습할 때 중간중간 오류가 나면 다 같이 해결하려고 머리 맞대는 과정이 빠졌지만,, 그 과정이 진짜 우리 스터디의 묘미였던 것 같아요ㅋㅋㅋ 덕분에 클라우드에 대한 자신감이 생겼습니다! 동아리 학술제 수상도 너무 뜻깊고 모두 열심히 노력한 덕분이에요! 열심히 참여해준 스터디원들 모두모두 감사합니다~!!



활동사진  
(3매 이상)





동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진



# 동아리 활동보고서

활 동 명	스터디 - 파이썬을 이용한 데이터 사이언스 & 시각화 스터디																																																																	
동아리 분과	학술분과	동아리명	KHUDA																																																															
활동시간	2024. 9. 25. ~ 2024. 11. 24. 매주 수요일 20:00 ~ 22:00	비고																																																																
활동목적	1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다. 2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다. 3. KHUDA 회칙 제4장 16조 [스터디] 목적: 스터디는 특정 주제에 대해 관심 있는 회원들이 자율적으로 모여, 해당 주제에 대해 학습 후 관련 경험을 쌓는다.																																																																	
활동내용 및 소감	<p>장소 : 공과대학 편의점 내 학습 공간</p> <p>활동내용 :</p> <p>본 스터디는 Python을 활용한 데이터 사이언스 및 시각화를 학습하며, 이를 통해 데이터 분석 역량을 향상시키는 것을 목표로 운영했습니다. [Python Data Science Handbook]을 기반으로 학습하며, 주요 라이브러리인 NumPy, Pandas, Matplotlib을 집중적으로 다루었습니다. 이를 통해 단순한 문법 습득을 넘어 실제 데이터 분석과 시각화에 활용할 수 있도록 했습니다.</p> <p>1) 운영 방식</p> <p>1. 주차별 학습 및 발제</p> <ul style="list-style-type: none"><li>- 매주 학습할 분량을 정한 후, 팀원 중 한 명이 발제를 맡아 학습 내용을 발표</li><li>- 발제자는 이론 정리뿐만 아니라, 최신 데이터 분석 동향이나 관련 이슈, 심화 발표 자료를 추가적으로 준비하고 공유</li></ul> <p>2. 질의응답 &amp; 토론</p> <ul style="list-style-type: none"><li>- 발제 후에는 질의응답 세션을 마련하여 스터디원들이 내용을 충분히 이해할 수 있도록 도움</li><li>- 단순한 강의 방식이 아니라, 함께 토론하며 개념을 확립하는 학습 방식 지향</li></ul> <p>3. 참여 독려 시스템 (벌금 제도 도입)</p> <ul style="list-style-type: none"><li>- 적극적인 참여를 유도하기 위해 사전 과제 지각 및 세션 지각 시 벌금 제도를 도입</li><li>- 이를 통해 스터디의 운영이 원활하게 이루어졌으며, 3주차 이후부터는 벌금이 쌓이지 않을 만큼 참여율을 높일 수 있었음</li></ul> <p>벌금 ...</p> <table><tr><th>Aa 이름</th><th>Σ 총</th><th># 1주차</th><th># 2주차</th><th># 3주차</th><th># 4주차</th><th># 5주차</th><th># 6주차</th><th># 7주차</th></tr><tr><td>김동혁</td><td>5,000</td><td>0</td><td>5,000</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>서지은</td><td>5,000</td><td>0</td><td>5,000</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>박정현</td><td>5,000</td><td>0</td><td>5,000</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>이두원</td><td>5,000</td><td>5,000</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>김형준</td><td>3,000</td><td>0</td><td>3,000</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>권서연</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>			Aa 이름	Σ 총	# 1주차	# 2주차	# 3주차	# 4주차	# 5주차	# 6주차	# 7주차	김동혁	5,000	0	5,000	0	0	0	0	0	서지은	5,000	0	5,000	0	0	0	0	0	박정현	5,000	0	5,000	0	0	0	0	0	이두원	5,000	5,000	0	0	0	0	0	0	김형준	3,000	0	3,000	0	0	0	0	0	권서연	0	0	0	0	0	0	0	0
Aa 이름	Σ 총	# 1주차	# 2주차	# 3주차	# 4주차	# 5주차	# 6주차	# 7주차																																																										
김동혁	5,000	0	5,000	0	0	0	0	0																																																										
서지은	5,000	0	5,000	0	0	0	0	0																																																										
박정현	5,000	0	5,000	0	0	0	0	0																																																										
이두원	5,000	5,000	0	0	0	0	0	0																																																										
김형준	3,000	0	3,000	0	0	0	0	0																																																										
권서연	0	0	0	0	0	0	0	0																																																										



## 2) 주차별 계획

Aa 주차	≡ 주요 개념	≡ 학습 내용	👥 기본 발제 준비자	👥 심화 발제 준비자
<b>1</b> 1주차	IPython Jupyter NumPy IDE	1장. Python - 파이썬에 날개를 달자 2장. NumPy 소개 2.1 파이썬의 데이터 타입 이해하기 2.2 NumPy 배열의 기초 2.3 NumPy 배열 연산: 유니버설 함수 2.4 집계: 최솟값, 최댓값, 그리고 그 사이의 모든 것 2.5 배열 연산: 브로드캐스팅	이 이두원	이 이두원
<b>2</b> 2주차	NumPy Pandas Series DataFrame	2장. NumPy 소개 2.6 비교, 마스크, 부울 로직 2.7 팬시 인덱싱 2.8 배열 정렬 2.9 구조화된 데이터: NumPy의 구조화된 배열 3장. Pandas로 데이터 가공하기 3.1 Pandas 설치 및 사용 3.2 Pandas 객체 소개 3.3 데이터 인덱싱과 선택 3.4 누락된 데이터 처리하기 3.5 계층적 인덱싱 3.6 데이터세트 결합: Concat과 Append 3.7 데이터세트 결합하기: 병합과 조인	정현 박 이 이두원	김 김동혁 서연 권
<b>3</b> 3주차	Pandas GroupBy pivot table	3장. Pandas로 데이터 가공하기 3.8 집계와 분류 3.9 피벗 테이블 3.10 벡터화된 문자열 연산 3.11 시계열 다루기 3.12 Pandas 시계열: 시간으로 인덱싱하기 3.13 고성능 Pandas: eval()과 query() 3.14 추가 자료	형 형준 김	서 서지은[학생](외국어대학 스
<b>4</b> 4주차	Matplotlib 히스토그램 밀도 플롯 등고선 플롯	4장. Matplotlib 사용법 4.1 일반적인 Matplotlib 사용법 4.2 하나 가격에 인터페이스 두 개 4.3 간단한 라인 플롯 4.4 간단한 산점도 4.5 오차 시각화하기 4.6 밀도 플롯과 등고선 플롯 4.7 히스토그램, 구간화, 밀도 4.8 플롯 범례 맞춤 변경하기	서 서지은[학생](외국어대학 스	김 김동혁
<b>5</b> 5주차	Matplotlib 다중 서브플롯 Seaborn	4장. Matplotlib 사용법 4.9 다중 서브플롯 4.10 텍스트와 주석 4.11 눈금 맞춤 변경하기 4.12 포맷 지시자와 위치 지시자 요약 4.13 Matplotlib 맞춤변경하기: 설정과 스타일시트 4.14 Matplotlib에서 3차원 플로팅하기 4.15 Basemap을 활용한 지리 데이터 4.16 Seaborn을 활용한 시각화 4.17 기타 자료	형 형준 김	김 김동혁
<b>6</b> 6주차	머신러닝 초모수 모델 검증 나이브 베이즈 다양체 가우스 혼합 모델	5장. 머신러닝 5.3 초모수와 모델 검증 5.4 특징 공학 5.5 심화 학습: 나이브 베이즈 분류 5.10 심화 학습: 다양체 학습 5.12 심화 학습: 가우스 혼합 모델 5.13 심화 학습: 커널 밀도 추정 5.14 응용: 안면 인식 파이프라인	김 김동혁	형 형준

### 3) 주차별 학습 내용 인증

Aa 이름	≡ 1주차	📎 2주차	📎 3주차	📎 4주차	📎 5주차	📎 6주차
김동혁	<a href="https://ddonghyeok0326.github.io/jekyll/update/2024/10/02/스터디-파이썬을-이용한-데이터-사이언스-8-시각화-스터디-week1.html">https://ddonghyeok0326.github.io/jekyll/update/2024/10/02/스터디-파이썬을-이용한-데이터-사이언스-8-시각화-스터디-week1.html</a>	<a href="https://ddonghye...">https://ddonghye...</a>	<a href="https://ddonghye...">https://ddonghye...</a>	<a href="https://ddonghye...">https://ddonghye...</a>	<a href="https://ddonghye...">https://ddonghye...</a>	<a href="https://ddonghye...">https://ddonghye...</a>
서지은	<a href="https://obvious-situation-664.notion.site/2-Numpy-1194505f659480a1aa64dc37cf62709f">https://obvious-situation-664.notion.site/2-Numpy-1194505f659480a1aa64dc37cf62709f</a>	<a href="https://obvious-sit...">https://obvious-sit...</a> <a href="http://obvious-sit...">http://obvious-sit...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://obvious-sit...">https://obvious-sit...</a>
박정현	<a href="https://github.com/Jeon-gDell/P_DataScience/blob/main/2-1.ipynb">https://github.com/Jeon-gDell/P_DataScience/blob/main/2-1.ipynb</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>	<a href="https://github.co...">https://github.co...</a>
김형준	<a href="https://github.com/caesar-kim/PythonDataScience/blob/main/first.md">https://github.com/caesar-kim/PythonDataScience/blob/main/first.md</a>	<a href="https://www.notio...">https://www.notio...</a>	<a href="https://discreet-dr...">https://discreet-dr...</a>	<a href="https://discreet-dr...">https://discreet-dr...</a>	<a href="https://discreet-dr...">https://discreet-dr...</a>	<a href="https://discreet-dr...">https://discreet-dr...</a>
이두원	<a href="https://muddy-clover-801.notion.site/1-Jupyter-1183374535a78013894bdf7cca9a4251?pvs=4">https://muddy-clover-801.notion.site/1-Jupyter-1183374535a78013894bdf7cca9a4251?pvs=4</a>	<a href="https://muddy-clo...">https://muddy-clo...</a> <a href="https://muddy-clo...">https://muddy-clo...</a>	<a href="https://muddy-clo...">https://muddy-clo...</a>	<a href="https://muddy-clo...">https://muddy-clo...</a>	<a href="https://muddy-clo...">https://muddy-clo...</a>	<a href="https://muddy-clo...5장 머신러닝.html">https://muddy-clo...5장 머신러닝.html</a>
권서연	<a href="https://blog.naver.com/ksy040407/223611603742">https://blog.naver.com/ksy040407/223611603742</a> , <a href="https://blog.naver.com/ksy040407/223612091776">https://blog.naver.com/ksy040407/223612091776</a>	<a href="https://blog.naver....">https://blog.naver....</a> <a href="https://blog.naver....">https://blog.naver....</a>	<a href="https://blog.naver....">https://blog.naver....</a>	<a href="https://blog.naver....">https://blog.naver....</a>	<a href="https://blog.naver....">https://blog.naver....</a>	<a href="https://blog.naver....">https://blog.naver....</a>

### 4) 스터디 주요 결과물

- **프로젝트 진행:** 최고의 항공사를 찾아라!
- 스터디에서 학습한 데이터 분석 및 시각화 기법을 활용하여 항공사별 정시 운항 성과를 분석하는 프로젝트를 진행했습니다.
- 단순한 분석에 그치지보다는 데이터에서 숨겨진 패턴을 발견하고 새로운 인사이트를 도출하는 것을 목표로 삼았습니다.
- **구체적인 프로젝트 목표:**
  1. 항공사별 데이터셋을 분석하여 어떤 항공사가 시간 약속을 잘 지키는지 평가
  2. 운항 우회율, 취소율, 지연율 등의 지표를 바탕으로 각 항공사의 시간 준수 성과 비교
  3. 시각화를 활용하여 항공사 운영 효율성과 고객 신뢰도를 한눈에 파악할 수 있도록 구성
- **스터디를 통해 쌓은 기술 스택**
  1. Python 기반 데이터 분석: NumPy, Pandas
  2. 데이터 시각화: Matplotlib, Seaborn
  3. 데이터 처리 및 전처리: Pandas를 활용한 데이터 정리 및 클리닝
  4. 인사이트 도출: 다양한 시각화 기법을 활용하여 의미 있는 패턴 발견

소감 :

이\*\*: 파이썬은 예전부터 사용해봤지만, 데이터 분석을 체계적으로 배운 건 이번이 처음이었어요. 단순히 코드를 따라 치는 게 아니라, 왜 이렇게 동작하는지 개념을 잡아가면서 배울 수 있어서 좋았습니다. 저는 무엇보다도 프로젝트를 진행하면서 실제 데이터에서 의미 있는 패턴을 찾아내는 과정이 흥미로웠어요! 앞으로도 꾸준히 데이터 분석을 공부해보고 싶어요.

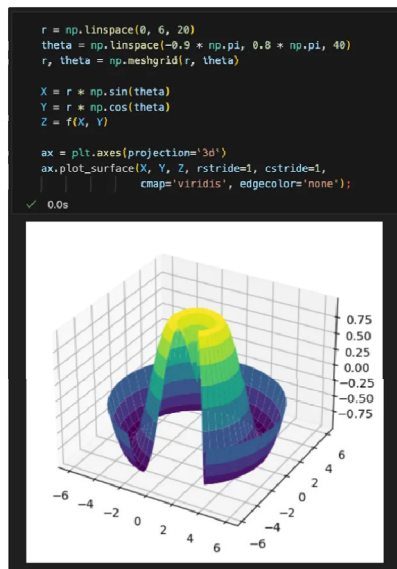
김\*\*: 저는 사실 KHUDA 신규 회원이기도 하고 처음엔 Pandas랑 Matplotlib이 낯설어서 걱정했는데, 직접 데이터를 분석하고 시각화하면서 점점 익숙해졌어요. 혼자 공부했다면 이해하기 어려웠을 것 같은 개념들도 발제자 발표를 듣고, 질문하면서 정리할 수 있어서 좋았어요. 특히 항공사 프로젝트 하면서 내가 만든 그래프가 실제 인사이트를 줄 수 있다는 게 신기했습니다.

권\*\*: 책만 보면 이해 안 되는 부분이 많았는데, 실습과 함께 배우니까 확실히 더 오래 기억에 남았어요. 주마다 벌금(?)이 있어서 강제적으로라도 복습하게 된 것도 도움이 됐고요ㅋㅋ. 프로젝트까지 진행하면서 단순한 이론 공부를 넘어서 실무 감각도 익힐 수 있어서 뿌듯했습니다. 다음엔 좀 더 복잡한 분석에도 도전해보고 싶어요!! 스터디원 모두모두 고생 많았습니다!!

#### 와이어프레임과 표면도

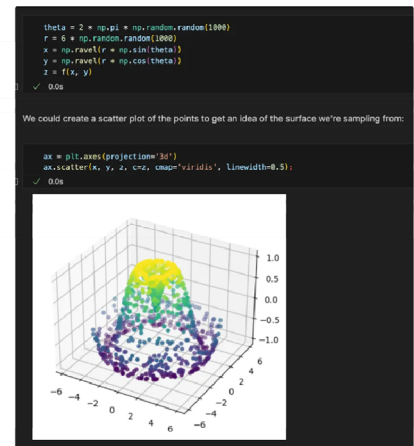
그리드 데이터 상에서 작동하는 또 다른 유형의 3차원 플롯으로 와이어프레임(wireframe)과 표면도가 있다. 이것들은 값의 그리드를 위해서 지정된 3차원 표면에 사영함으로써 3차원 형태의 결과를 시각화할 수 있다.

시각화하는 함수에 슬라이스를 줄 수 있는 부분 극좌표 그리드(partial polar grid)를 만들어보자.



#### 표면 삼각측량법

직교 좌표계나 극좌표 그리드에서 균일하게 추출하는 대신 무작위로 추출하기



활동사진  
(3매 이상)

## 심화 학습: 다양체 학습

다양체 학습은 비선형 관계를 설명하는 비지도 학습 방법이다.

고차원 데이터에서 저차원 다양체를 찾아 데이터의 핵심적인 관계를 유지하려고 한다.

주요 기법으로는 다차원 척도법(MDS), 국부 선형 임베딩(LLE), 등거리 사상(isomap) 등이 있다.

다양체 학습으로서 MDS는 데이터의 거리 행렬을 이용해 고차원 데이터를 저차원 공간으로 사영하여 본질적인 구조를 유지한다.

```
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns; sns.set()
import numpy as np
```

✓ 2.2s Python

```
from mpl_toolkits import mplot3d
ax = plt.axes(projection='3d')
ax.scatter(X3[:, 0], X3[:, 1], X3[:, 2],
          **colorize)
ax.view_init(elev=70, azim=50)
```

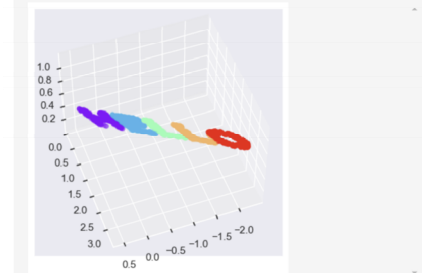
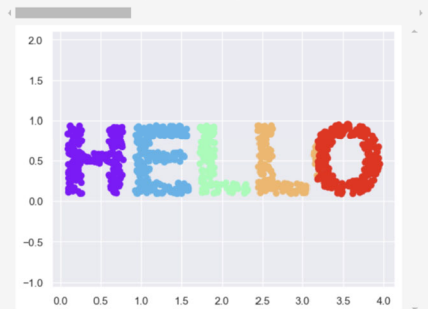
✓ 0.2s Python

## 다양체 학습: 'HELLO'

```
X = make_hello(1000)
colorize = dict(c=X[:, 0], cmap=plt.cm.get_cmap('rainbow', 5))
plt.scatter(X[:, 0], X[:, 1], **colorize)
plt.axis('equal');
```

✓ 0.3s Python

```
C:\Users\kysm\Anaconda3\envs\envkernel_45336\lib\site-packages\matplotlib\cm.py:2: MatplotlibDeprecationWarning:
colorize = dict(c=X[:, 0], cmap=plt.cm.get_cmap('rainbow', 5))
```







KHUDA

이벤트

정보

# 공지

# tech-talk

# 개발-행사

# 개발-대회

# 개발-트렌드

음성 채널

운영진 회의방

자율 회의방 1

6기 YB 김동혁 라이브

5기 YB 이두원 / NLP

6기 OB 권서연

6기 OB 서지은

6기 YB 박정현

김형준

자율 회의방 2

자율 회의방 3

모각공 1

모각공 2

모각공 3

? 질문방

음성 연결됨

자율 회의방 1 / KHUDA

회원

관리자

자율 회의방 1

6기 YB 김동혁님의 화면

720p 30FPS 라이브

다양체 학습으로로서의 MDS

- MDS 추정기에 3차원 데이터를 입력하고 거리 행렬을 계산한 다음, 이 거리 행렬에 대한 최적의 2차원 임베딩을 결정하도록 요청
- 결과를 원본 데이터의 표현을 복구

```
model = MDS(n_components=2, random_state=1)
out3 = model.fit_transform(X3)
plt.scatter(out3[:, 0], out3[:, 1], **colorize)
plt.axis('equal');
```

6기 YB 김동혁

김형준

5기 YB 이...

6기 YB 박...

6기 OB 권...

6기 OB 서...



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

정유진

# 동아리 활동보고서

활 동 명	스터디 - Apache Kafka 애플리케이션 프로그래밍 with JAVA		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 9. 9. ~ 2024. 11. 24. 매주 목요일 21:30 ~ 23:30	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다.</p> <p>2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>3. KHUDA 회칙 제4장 16조 [스터디] 목적: 스터디는 특정 주제에 대해 관심 있는 회원들이 자율적으로 모여, 해당 주제에 대해 학습 후 관련 경험을 쌓는다.</p>		
활동내용 및 소감	<p>장소 : 중앙도서관 스터디룸, 공과대학 편의점 내 학습 공간</p> <p>활동내용 :</p> <p>본 스터디는 [Apache Kafka Application Programming with JAVA] 책을 기반으로 진행되었으며, Kafka의 핵심 개념을 빠르게 익힌 후 실습을 병행하여 <u>이론과 실무를 연결하는 경험을 쌓는 것을 목표로 하였습니다.</u> 단순히 개념을 학습하는 데 그치지 않고, <u>peer review</u> 및 세미나 발표를 통해 심화 학습과 실질적인 적용 방법을 고민하는 과정을 거쳤습니다.</p> <p>1) 스터디 참여 태도</p> <p>이 스터디는 단순한 학습이 아니라 서로에게 배움을 주고받는 환경을 만드는 것이 목표였음. 따라서 스터디원들에게 다음과 같은 참여 태도를 강조하고 서로 약속했음.</p> <ul style="list-style-type: none"> <li>- 모르는 것을 부끄러워하지 않고 적극적으로 질문하자!</li> <li>- 본인이 아는 것과 모르는 것을 팀원들과 함께 나누자!</li> </ul> <p>2) 운영 방식</p> <p>1. 주차별 학습 &amp; 발표</p> <ul style="list-style-type: none"> <li>- 모든 스터디원은 매주 정해진 분량을 학습해 오며, 1~2명이 발표자로 선정되어 주요 개념을 요약하여 발표(PDF, PPT 권장).</li> <li>- 단순한 내용 전달이 아니라, 핵심 내용을 요약하여 정리하고, 직접 고민한 내용을 함께 나누는 방식으로 진행.</li> <li>- 발표 시간은 10~15분으로 제한하여, 효율적으로 내용을 정리하고 전달하는 능력을 기를 수 있도록 함.</li> </ul> <p>2. 세미나 리뷰 &amp; 기술 블로그 공유</p> <ul style="list-style-type: none"> <li>- 매주 한 명이 해당 주제와 관련된 테크 블로그, 컨퍼런스 영상 등을 찾아 세미나 발표 진행.</li> <li>- 단순한 전달이 아니라, 팀원들이 이해하기 쉽게 정리하고, 실제 사례와 비교하여 설명하는 방식으로 운영.</li> </ul>		

- Kafka가 실제로 어떻게 활용되고 있는지, 실무에서 어떤 문제를 해결하는지를 탐구하는 과정.

### 3. 실습 중심 & 협업 해결 지향

- Kafka는 실습이 필수적인 기술이므로, 모든 스터디원이 함께 실습을 진행.
- 누군가 실습에 문제가 생기면 해결할 수 없는 H/W 이슈가 아닌 이상, 모두가 함께 원인을 분석하고 해결하는 방식으로 진행.
- 이를 통해 단순한 개별 학습이 아니라, 문제 해결 역량과 협업 능력을 함께 키울 수 있는 스터디가 됨.

### 3) 스터디 계획 및 진행 기록 (모두 가져오기에는 내용이 많아 일부만 가져왔습니다)

Aa 차시	≡ 비고	≡ 분량	≡ 세미나 링크	교재 reviewer	세미나 reviewer
1 차시		<p>[1장] 들어가며</p> <p>1.1 카프카의 탄생</p> <p>1.2 빅데이터 파이프라인에서 카프카의 역할</p> <p>1.3 데이터 레이크 아키텍처와 카프카의 미래</p> <p>1.4 정리</p> <p>[2장] 카프카 빠르게 시작해보기</p> <p>2.1 실습용 카프카 브로커 설치</p> <p>2.1.1 AWS EC2 인스턴스 발급 및 보안 설정</p> <p>2.1.2 인스턴스에 접속하기</p> <p>_____ssh 명령어로 접속하기</p> <p>_____putty로 접속하기</p> <p>2.1.3 인스턴스에 자바 설치</p> <p>2.1.4 주키퍼-카프카 브로커 실행</p> <p>_____카프카 브로커 힙 메모리 설정</p> <p>_____카프카 브로커 실행 옵션 설정</p> <p>_____주키퍼 실행</p> <p>_____카프카 브로커 실행 및 로그 확인</p> <p>2.1.5 로컬 컴퓨터에서 카프카와 통신 확인</p> <p>_____테스트 편의를 위한 hosts 설정</p> <p>2.2 카프카 커맨드 라인 툴</p> <p>2.2.1 kafka-topics.sh</p> <p>_____토픽 생성</p> <p>_____토픽 리스트 조회</p> <p>_____토픽 상세 조회</p> <p>_____토픽 옵션 수정</p> <p>2.2.2 kafka-console-producer.sh</p> <p>2.2.3 kafka-console-consumer.sh</p> <p>2.2.4 kafka-consumer-groups.sh</p> <p>2.2.5 kafka-verifiable-producer, consumer.sh</p> <p>2.2.6 kafka-delete-records.sh</p> <p>2.3 정리</p>	<p><b>담당영역</b> [Apache Kafka 도입기]</p> <p><a href="https://medium.com/mildan-g/apache-kafka-도입기-3451787d97a8">https://medium.com/mildan-g/apache-kafka-도입기-3451787d97a8</a></p> <p><b>여기어때</b> [Apache Kafka를 사용하여 EDA 적용하기]</p> <p><a href="https://techblog.gccompany.co.kr/apache-kafka를-사용하여-eda-적용하기-bf263c79efd0">https://techblog.gccompany.co.kr/apache-kafka를-사용하여-eda-적용하기-bf263c79efd0</a></p>	유혜지	김건형
3 차시		<p>[3장] 카프카 기본 개념 설명 - 2</p> <p>3.5 카프카 스트림즈</p> <p>3.5.1 스트림즈DSL</p> <p>_____KStream</p> <p>_____KTable</p> <p>_____GlobalKTable</p> <p>_____스트림즈DSL 주요 옵션</p> <p>_____스트림즈DSL - stream(), to()</p> <p>_____스트림즈DSL - filter()</p> <p>_____스트림즈DSL - KTable과 KStream을 join()</p> <p>_____스트림즈DSL - GlobalKTable과 KStream을 join()</p> <p>3.5.2 프로세서 API</p> <p>3.6 카프카 커넥트</p> <p>_____커넥트를 실행하는 방법</p> <p>_____단일 모드 커넥트</p> <p>_____분산 모드 커넥트</p> <p>3.6.1 소스 커넥터</p> <p>_____파일 소스 커넥터 구현</p> <p>3.6.2 싱크 커넥터</p> <p>_____파일 싱크 커넥터 구현</p> <p>3.7 카프카 미러메이커2</p> <p>3.7.1 미러메이커2를 활용한 지리적 복제(Geo-Replication)</p> <p>3.8 정리</p>	<p><b>카카오</b> [카카오   카프카 커넥트를 데이터 파이프라인으로 사용하는 이유?]</p> <p><a href="https://tech.kakao.com/posts/564">https://tech.kakao.com/posts/564</a></p>	유혜지 김건형	이 이재호



## Questions: more than two questions per person

### 윤소은

- 클러스터의 브로커 중에서 컨트롤러 역할과 코디네이터 역할을 하는 브로커가 있어야하는데, 이 역할이 중복이 될 수 있는지 ?

가능하다!

- 브로커로 데이터 보내기 위해 ProducerRecord를 생성할 때, `ProducerRecord<String, String> record = new ProducerRecord<>(TOPIC_NAME, messageValue);`에서 `<String, String>`은 TOPIC\_NAME, messageValue를 선언하는건지, 메시지의 key, value값 타입을 선언하는건지?

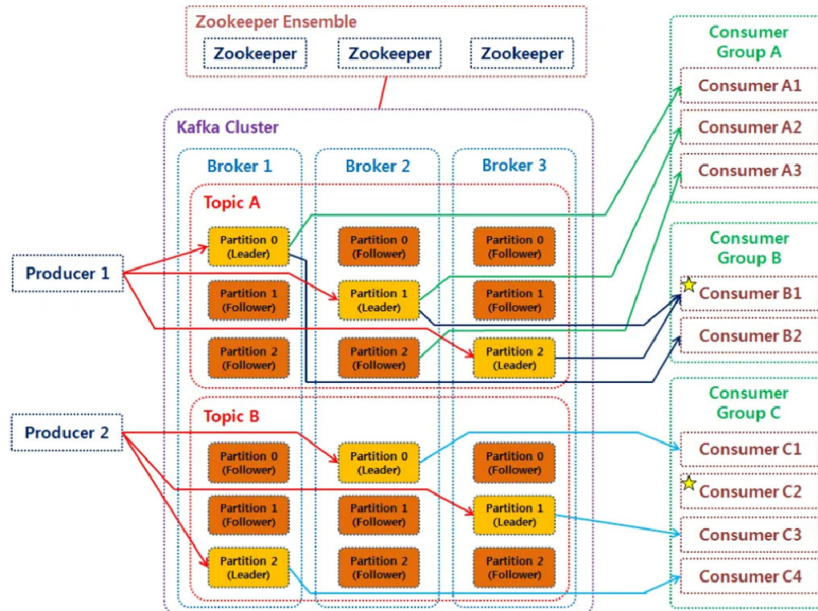
### Constructors

#### Constructor and Description

**ProducerRecord(String topic, Integer partition, K key, V value)**  
Creates a record to be sent to a specified topic and partition

### 황중훈

- 하나의 브로커 안에 여러 개의 리더 파티션이 존재할 수 있는가?



있다! 브로커의 Topic A의 파티션 개수가 3개, 복제 개수(=Broker의 수)가 3개일 때, 브로커가 3개 → 2개로 줄었  
다면 Leader를 다른 브로커로 넘겨줘야 함 (follower가 leader로 활성화). 복제 개수는 자동으로 2개로 줄게 되고,  
이에 따라 Broker1 또는 Broker2 둘 중 하나는 Leader가 두 개가 될 것임.  
조금 더 쉽게는 처음부터 브로커 1개만 있다면 그 브로커에 있는 모든 파티션이 리더일 것

## 4) 스터디 결과물

- 프로젝트 진행: 실시간 서울시 상권 데이터 분석

- 프로젝트 설명:

1. Apache Kafka 기반 데이터 용도에 따른 세부 API 제공을 목표로 프로젝트 진행
2. 서울 열린 데이터광장에서 제공하는 '서울시 실시간 도시 데이터'를 활용하여 Kafka 기반 데이터 스트리밍을 설계

- **프로젝트 목표:**

1. Kafka Streams 및 Kafka Connector를 활용하여 서울시 실시간 도시 데이터를 효과적으로 수집 및 가공
2. 실시간 데이터 스트리밍 환경에서 적절한 소스에서 데이터를 수집하고, Kafka 토픽을 분기 및 적재하는 과정을 직접 구현
3. Kafka의 높은 확장성과 가용성을 직접 체험하고, 실무에서 ETL(Data Extraction, Transformation, Load) 파이프라인 구축에 적용할 수 있는 역량을 갖추는 것

**소감 :**

**이\*\*:** Kafka가 어렵다는 건 들었지만, 설치부터 이렇게 삼질할 줄은 몰랐다. 처음엔 뭐가 뭔지 하나도 모르겠어서 환경 설정하는 데만 몇 시간을 썼는데.. 다행히 스터디원들이랑 같이 해결하면서 점점 익숙해졌다. Producer-Consumer가 데이터를 주고받는 과정을 직접 구현해보니까 그동안 개념으로만 이해했던 메시지 큐 시스템이 머릿속에서 확실히 정리됐다. 이제는 Kafka 클러스터를 직접 띄우고, 토픽을 관리하는 게 어색하지 않다.

**윤\*\*:** 책만 읽을 때는 그냥 대충 이해한 것 같았는데, 발표하려고 내용을 정리하다 보니까 모르는 게 엄청 많다는 걸 깨달았어요. **단순히 문장을 읽는 게 아니라, 다른 사람한테 설명할 수 있을 정도로 이해해야 하니까 더 깊이 공부하게 됐어요.** 책에서 Kafka Streams랑 Connector 부분은 개념만 보면 어려운데, 진짜 예제 코드로 실습하면서 '아하! 이래서 이렇게 쓰는구나' 하고 감이 왔어요. 발표 준비 과정은 힘들었지만 덕분에 다른 팀원들한테도 좋은 피드백을 받을 수 있어서 뿌듯했어요!

**홍\*\*:** 다른 프로젝트를 진행하면서 데이터 파이프라인을 다루긴 했지만 Kafka는 처음 접해봤다. 기존에는 데이터 이동을 주로 배치(batch) 방식으로 처리했었는데, Kafka를 공부하면서 실시간 스트리밍이 왜 중요한지 확실히 이해하게 됐다. 프로젝트에서 서울시 실시간 데이터를 다루니까 실무에서 Kafka가 어떻게 활용될 수 있을지 감이 오는 것 같다. 앞으로 데이터 엔지니어링 업무에서 Kafka를 적극적으로 활용할 수 있을 것만 같은 자신감이 넘친다!!!!

활동사진  
(3매 이상)







동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

(정유진)



## 동아리 활동보고서

활 동 명	스터디 - 도커 & 쿠버네티스 스터디		
동아리 분과	학술분과	동아리명	KHUDA
활동시간	2024. 7. 23. ~ 2024. 9. 3. 매주 화요일 19:00 ~ 21:00	비고	
활동목적	<p>1. KHUDA 회칙 제1장 제3조 [목적] 기본 커리큘럼 및 스터디를 통한 데이터 분석, 머신러닝, 딥러닝의 기초 지식을 함양한다.</p> <p>2. KHUDA 회칙 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p> <p>3. KHUDA 회칙 제4장 16조 [스터디] 목적: 스터디는 특정 주제에 대해 관심 있는 회원들이 자율적으로 모여, 해당 주제에 대해 학습 후 관련 경험을 쌓는다.</p>		
활동내용 및 소감	<p>장소 : 중앙도서관 스터디룸 또는 정문 근처 카페</p> <p>활동내용 :</p> <p>본 스터디는 기업에서 필수적으로 활용되는 도커(Docker)와 쿠버네티스(Kubernetes)의 개념과 실무 활용법을 학습하는 것을 목표로 개설했습니다. 컨테이너 기술의 핵심 개념을 익히고, 실습을 통해 직접 환경을 구축하며 배포 자동화 및 클러스터 관리의 원리를 이해하는 과정으로 진행되었습니다.</p> <p>1) 운영 방식</p> <ul style="list-style-type: none"> <li>1. 교재 기반 학습 &amp; 발표 <ul style="list-style-type: none"> <li>- 《한 권으로 배우는 도커 &amp; 쿠버네티스》 교재를 기반으로 스터디를 진행했습니다.</li> <li>- 매주 정해진 학습 범위를 스터디원 개별적으로 학습한 후, 발표자가 핵심 개념을 정리하여 발표했습니다.</li> <li>- 발표 후에는 스터디원들이 공부하면서 궁금했던 점을 질문하고 토론하는 방식으로 진행하여, 놓치는 부분 없이 개념을 깊이 이해할 수 있도록 했습니다.</li> </ul> </li> <li>2. 실습 중심 학습 &amp; 상호 피드백 지향 <ul style="list-style-type: none"> <li>- 도커와 쿠버네티스는 환경 설정이 복잡하여 실습 중 오류가 자주 발생하므로, 각자 실습을 진행한 후 발생한 문제를 공유하고 함께 해결하는 방식으로 진행했습니다.</li> <li>- 실습 진행 시, 해당 주차 발표자는 이론적인 개념을 간단히 정리하고, 실습 과정에서 얻은 인사이트를 공유했습니다.</li> <li>- VM을 활용한 리눅스 환경에서 실습을 진행하여, 환경 차이에서 오는 문제를 최소화하고 원활한 피드백이 가능하도록 했습니다.</li> </ul> </li> </ul> <p>2) 주차별 계획 및 기록</p>		

Aa 주차	주차주요 개념	학습목표	발제자	심화 발...
<b>1</b> 1주차	환경 설정	1~3장 도커의 실습 환경 구축 및 기초 지식 학습	이 이재호	
<b>2</b> 2주차	도커	4장 도커의 기초에 대해 학습	이소연	
<b>3</b> 3주차	도커 django Flask	5~6장 도커를 활용한 실습	김민아	박상영
<b>4</b> 4주차	환경 설정	7~8장 쿠버네티스의 실습 환경 구축 및 기초 지식 학습	장 장서연	장 장서연
<b>5</b> 5주차	쿠버네티스	9장 쿠버네티스의 기초에 대해 학습	박상영	김민아
<b>6</b> 6주차	인그레스	10장 쿠버네티스를 활용해 웹 서비스 배포	이 이재호	이소연
<b>7</b> 7주차	github 프로메테우스 CI/CD	11~12장 쿠버네티스 클러스터 관리 및 소스코드 관리	류 류여진	이 이재호

#### 주차별 학습 내용

Aa 이름	1주차	2주차	3주차	4주차	5주차	6주차	7주차
김민아	<a href="https://magenta-...">https://magenta-...</a>	<a href="https://magenta-...">https://magenta-...</a>	<a href="https://magenta-...">https://magenta-...</a>	<a href="https://magenta-...">https://magenta-...</a>	<a href="http://magenta-m...">http://magenta-m...</a> <a href="https://magenta-...">https://magenta-...</a>		<a href="https://magenta-...">https://magenta-...</a>
다니아	<a href="https://daniazie.git...">https://daniazie.git...</a> <a href="https://daniazie.git...">https://daniazie.git...</a> <a href="https://daniazie.git...">https://daniazie.git...</a>	<a href="https://daniazie.git...">https://daniazie.git...</a>	<a href="https://daniazie.git...">https://daniazie.git...</a> <a href="https://daniazie.git...">https://daniazie.git...</a>	<a href="https://daniazie.git...">https://daniazie.git...</a> <a href="https://daniazie.git...">https://daniazie.git...</a>		<a href="https://daniazie.git...">https://daniazie.git...</a>	<a href="https://daniazie.git...">https://daniazie.git...</a>
류여진	<a href="https://scientific-e...">https://scientific-e...</a>	<a href="https://scientific-e...">https://scientific-e...</a>	<a href="https://scientific-e...">https://scientific-e...</a>	<a href="https://scientific-e...">https://scientific-e...</a>			<a href="https://scientific-e...">https://scientific-e...</a> <a href="https://scientific-e...">https://scientific-e...</a>
박상영	<a href="https://sypdevlog-...">https://sypdevlog-...</a>	<a href="https://sypdevlog-...">https://sypdevlog-...</a>	<a href="https://sypdevlog-...">https://sypdevlog-...</a> <a href="https://sypdevlog-...">https://sypdevlog-...</a> <a href="https://sypdevlog-...">https://sypdevlog-...</a>	<a href="https://sypdevlog-...">https://sypdevlog-...</a> <a href="https://sypdevlog-...">https://sypdevlog-...</a>	<a href="https://sypdevlog-...">https://sypdevlog-...</a>		
이소연	<a href="https://velog.io/@...">https://velog.io/@...</a>	<a href="https://velog.io/@...">https://velog.io/@...</a>	<a href="https://velog.io/@...">https://velog.io/@...</a>	<a href="https://velog.io/@...">https://velog.io/@...</a>		<a href="https://velog.io/@...">https://velog.io/@...</a>	
이재호	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a>	<a href="https://leedominic-...">https://leedominic-...</a> <a href="https://leedominic-...">https://leedominic-...</a>
장서연	<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>	<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>	<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>	<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>	<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>		<a href="https://onyx-eclip-...">https://onyx-eclip-...</a>

#### 소감 :

**이\*\*:** 평소에 도커라는 단어는 많이 들어봤는데 실제로 써본 건 이번이 처음이었다. 처음엔 가상 머신이랑 뭐가 다른 건지도 몰랐는데 실습을 진행해보면서 점점 개념이 잡혔다. 컨테이너 이미지를 만들고 배포해보는 과정이 제일 흥미로웠다. 스터디 참여하면서 실제 운영 환경에서 도커가 어떻게 쓰이는지 확실하게 이해했고, 다음 개인 프로젝트에도 도커를 적용해볼 생각이다.

**류\*\*:** 솔직히 실습할 때 오류가 너무 많이 나서 멘붕 온 적도 많았어요ㅠㅠ 근데 다 같이 오류

활동사진  
(3매 이상)

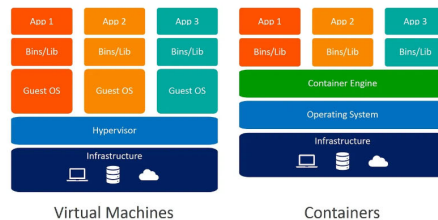
를 해결하면서 배우는 게 진짜 도움이 많이 됐어요!!! 혼자 공부했으면 '왜 안 되지?' 하다가 포기했을 텐데, 스터디에서 문제를 공유하고 같이 해결하니까 더 깊이 이해할 수 있었어요~~ 쿠버네티스에서 Pod가 제대로 동작하지 않을 때 로그를 확인하는 방법이나 네트워크 설정을 디버깅하는 과정에서 배우는 것이 많았어요. 다들 스터디하는 동안 고생 많았어요!!

[KHUDA] 도커 스터디 1주차

### 2.2.1 도커의 정의

- 도커(Docker)는 컨테이너 기반 가상화 플랫폼으로, 응용 프로그램과 그 종속성을 격리된 환경인 컨테이너로 패키징하여 실행하는 기술
  - 응용 프로그램을 서로 다른 환경에서도 일관되게 실행 가능
  - 개발 환경과 운영 환경 사이의 차이로 인한 문제를 줄일 수 있음
  - 도커 컨테이너는 가볍고 빠르게 확장성이 좋아서 개발 및 배포 프로세스를 간소화하는 데 사용
- + 컨테이너: 가상화 기술을 이용하여 어플리케이션과 개발 환경을 격리된 공간에서 실행하는 단위

- Virtual Machines와 Docker
  - 컨테이너 기술이 아닌 기존의 가상화 방식은 주로 OS를 가상화
  - VMware, VirtualBox와 같은 가상머신은 호스트 OS 위에 게스트 OS 전체를 가상화하여 사용하는 방식으로 도커의 컨테이너 방식과는 큰 차이



- Virtual Machines
  - 가상머신은 호스트 운영체제 위에 가상화된 하드웨어 계층을 생성하고, 각 가상 머신은 독립된 운영체제, 커널, 드라이버 등을 가짐
  - 이로 인해 무겁고 높은 자원 소비가 필요
  - 운영체제의 부팅 과정이 필요하므로 시간이 오래 걸림
  - 독립된 운영체제를 가지므로 메모리, 디스크 공간 등 자원을 많이 소비

### 도커를 활용한 django 실행

#### 도커 호스트에 django 프로젝트 생성

- ch05 폴더에서 ex02라는 디렉터리를 생성하여 들어간다

```

ania@server01:~/work/ch05/ex01$ cd ..
ania@server01:~/work/ch05$ mkdir ex02
ania@server01:~/work/ch05$ cd ex02
ania@server01:~/work/ch05/ex02$
    
```

- 새로운 django 프로젝트를 생성하고 파일 목록 확인

```

ania@server01:~/work/ch05/ex02$ pyenv activate py3_11_6
(py3_11_6) ania@server01:~/work/ch05/ex02$ django-admin startproject myapp
(py3_11_6) ania@server01:~/work/ch05/ex02$ ls
myapp
(py3_11_6) ania@server01:~/work/ch05/ex02$ tree ./
.
├── myapp
│   ├── manage.py
│   └── myapp
│       ├── asgi.py
│       ├── __init__.py
│       ├── settings.py
│       ├── urls.py
│       └── wsgi.py
└── 3 directories, 6 files
    
```

- vim settings.py 를 입력하고 settings.py를 아래와 같이 수정한다

```

# SECURITY WARNING: don't run with debug turned on in production!
DEBUG = True

ALLOWED_HOSTS = ['*']
    
```

- django 프로젝트를 실행하기 전에 데이터베이스 변경사항을 적용한다

```

(py3_11_6) ania@server01:~/work/ch05/ex02/myapp$ cd ..
l(py3_11_6) ania@server01:~/work/ch05/ex02/myapp$ ls
manage.py myapp
    
```

#### Contents

- 추가 학습 환경 구축
- YAML 기초
- 도커를 활용한 django 실행**
- 도커 호스트에 django 프로...
- django 이미지 빌드
- django 컨테이너 실행
- Nginx, django 연동 후 실행
- Nginx, django, PostgreSQL 컨...
- Nginx, django와 로컬 Postgre...
- 도커 컴포즈를 활용한 컨테이...



## 5주차 - 스테이트풀셋



책에서 설명하는 스테이트풀셋(StatefulSet)

- 1) 파드들을 관리한다는 점에서 디플로이먼트와 유사
- 2) 그러나 디플로이먼트와 다르게 각 파드의 독자성을 유지
- 3) 따라서 파드들이 동일한 스펙으로 생성되었다고 해도 각 파드들을 서로 교체해서 사용할 수 없음 (바꿔 말하면, 디플로이먼트는 각 파드들이 서로 대체 가능)
- 4) 각 파드는 영구적인 식별자를 가지며 이를 식별하기 위해 헤드리스 서비스를 사용한다

### 1) 파드들을 관리한다는 점에서 디플로이먼트와 유사



StatefulSets

공식 사이트에서 제공하는 설명은 다음과 같다.

- StatefulSet is the workload API object used to manage **stateful applications**.
- 파드 집합의 배포 및 확장을 관리하고, 파드의 순서 및 **고유성**에 대한 보증을 제공한다
- Deployment와 마찬가지로, 동일한 컨테이너 사양을 기반으로 하는 **파드를 관리한다**

등등..

- **컨테이너들의 집합인 파드**를 관리 및 유지하기 위한 컨트롤러인 **레플리카셋**, 그리고 **레플리카셋을 관리하는 컨트롤러(\*)**이자 **애플리케이션의 배포와 스케일링을 관리하는 디플로이먼트**와 유사하게, 스테이트풀셋도 역시 파드들을 관리한다.
- 스테이트풀셋 생성 시 디플로이먼트처럼 파드의 수를 설정하고, 서비스와 연동하여 파드를 생성할 수 있다.
- 스테이트풀셋에서는 **안정 및 지속**을 중요시 여겨, 파드 삭제시 랜덤한 파드가 삭제되는 것이 아닌, 최신의 파드를 삭제한다. ⇒ 애플리케이션의 **안정성과 순차적인 배포**가 필요할 때 활용한다.

\* 디플로이먼트는 레플리카셋에게 revision을 남겨 롤백을 가능하게 하는 등의 기능을 제공한다

동아리활동과 관련하여 상기 결과보고서를 확인합니다.

※본 활동보고서가 '활동우수지원금' 지급에 평가 기준으로 활용됨에 동의합니다.

대표학생 : 정유진

392