

동아리 활동보고서

활동명	기초세션		
동아리 분과	학술 분과	동아리명	KHUDA
활동시간	2025.07.22 ~ 08.29	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 강의실</p> <p>활동내용 : 스터디는 머신러닝 입문자를 포함한 동아리원들이 함께 참여하여, 기본 개념부터 단계적으로 학습할 수 있도록 구성되었다. 주요 학습 내용은 다음과 같다.</p> <ul style="list-style-type: none"> 머신러닝의 개요 및 전체 파이프라인 이해 지도학습과 비지도학습의 개념 및 차이 선형 회귀, 로지스틱 회귀 등 기본적인 모델의 원리 모델 학습 과정, 성능 평가 방법 과적합 및 일반화 개념 간단한 데이터 전처리 및 모델 적용 사례 <p>각 팀은 매주 정해진 분량의 교재 내용을 사전에 학습한 후, 핵심 내용을 정리하여 발표 자료를 준비하였다. 발표에서는 이론 설명뿐만 아니라 예제 코드, 그림 등을 활용하여 이해를 돕고자 하였다.</p> <p>김 XX 머신러닝을 혼자 공부할 때는 개념이 막연하게 느껴졌는데, 팀별로 발표하고</p>		

토론하는 과정을 통해 훨씬 명확하게 이해할 수 있었다. 특히 다른 팀의 설명을 들으며 놓치고 있던 부분을 다시 정리할 수 있어 도움이 되었다.

이 XX

핸즈온 머신러닝 교재를 처음부터 끝까지 꾸준히 읽는 것이 쉽지 않았는데, 스터디 형식으로 진행되니 책임감을 가지고 참여할 수 있었다. 이론뿐만 아니라 실제 적용 흐름을 이해하는 데 큰 도움이 된 활동이었다.

박 XX

발표를 준비하면서 개념을 스스로 설명할 수 있을 정도로 정리해야 했기 때문에 학습 효과가 높았다. 머신러닝 전반에 대한 기초를 다질 수 있는 의미 있는 시간이었다.

활동사진

(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬 박(윤찬)

동아리 활동보고서

활동명	데이터톤		
동아리 분과	학술 분과	동아리명	KHUDA
활동시간	2025. 11월 22일	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 제1장 제3조 [목적] 회원들 간의 상호 협력과 지식 공유를 촉진하여 네트워크 형성과 교류의 장을 마련한다.</p>		
활동내용 및 소감	<p>장소 : 전자정보대학 205호</p> <p>1) 활동 개요</p> <p>본 데이터톤은 KHUDA 동아리 주관으로 진행된 데이터 분석 중심 해커톤 형태의 활동으로, KHUDA 부원뿐만 아니라 다양한 학교의 참가자들이 함께 참여하여 진행되었다. 참가자들은 팀을 구성하여 주어진 주제에 맞는 데이터 분석과 서비스 아이디어 도출을 수행하였으며, 데이터 기반 인사이트를 중심으로 결과물을 발표하였다.</p> <p>2) 주제 및 과제 내용</p> <p>데이터톤의 주제는 **「연령대별 x 의식주」**를 중심으로, 한국인의 라이프스타일에 맞는 데이터 기반 서비스 또는 웹/앱 아이디어를 기획·분석하는 것이었다. 각 팀은 연령대와 의식주 요소를 자유롭게 조합하여 분석 대상을 설정하고, 해당 타겟에 적합한 서비스를 설계하였다.</p> <p>주요 과제는 다음과 같다.</p> <p>연령대와 의식주 조합 선정 및 분석 방향 설정</p>		

공개 데이터 또는 외부 데이터 API를 활용한 데이터 분석

데이터 분석 결과를 통해 도출한 인사이트를 기반으로 서비스 기획

머신러닝 또는 데이터 분석 기법을 활용한 기능 설계

실제 사용을 가정한 웹/앱 형태의 결과물 구상

3) 활동 진행 방식

데이터톤은 정해진 일정에 따라 단계적으로 진행되었다. 참가자들은 팀별로 데이터 탐색과 분석을 수행하고, 분석 결과를 토대로 서비스의 핵심 기능과 차별점을 구체화하였다. 이 과정에서 단순한 아이디어 제안이 아닌, 데이터 분석 결과에 근거한 문제 정의와 해결 방안을 제시하는 데 중점을 두었다.

또한 모델 예측, 분류, 군집화 등 머신러닝 기법이나 데이터 시각화 결과를 활용하여 서비스의 타당성을 설명하도록 하였으며, 최종적으로는 실제 사용 가능한 웹 또는 앱 형태의 결과물을 목표로 하여 발표를 진행하였다.

4) 활동 성과

본 데이터톤을 통해 참가자들은 제한된 시간 안에 데이터를 분석하고 인사이트를 도출하는 경험을 쌓을 수 있었다. 특히 다양한 학교의 참가자들과 협업하며 서로 다른 관점과 접근 방식을 공유할 수 있었고, 데이터 분석이 서비스 기획과 어떻게 연결되는지에 대한 실질적인 이해를 높일 수 있었다.

9. 데이터톤 참여자 소감

김XX

다양한 학교의 참가자들과 한 팀으로 프로젝트를 진행하며 새로운 시각을 많이 배울 수 있었다.

데이터 분석 결과를 바탕으로 서비스 아이디어를 구체화하는 과정이 인상 깊었다.

이XX

단기간에 데이터 분석부터 서비스 기획까지 모두 경험할 수 있어 밀도 높은 활동이었다. 실제로 활용 가능한 결과물을 만드는 것이 목표라서 더욱 집중해서 참여할 수 있었다.

김 XX

기존에 배웠던 머신러닝과 데이터 분석 기법이 실전에서 어떻게 사용되는지 체감할 수 있었다.

협업과 발표

활동사진
(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬 박윤찬

동아리 활동보고서

활동명	Data Engineering – 심화트랙		
동아리 분과	학술	동아리명	KHUDA
활동시간	2025. 09. 01 ~ 2026. 01. 07	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 : 매 주차별 상이</p> <p>1) 활동 개요</p> <p>Data Engineering 트랙은 한 학기 동안 데이터 엔지니어링 전반의 흐름을 이해하고, 실제 데이터 파이프라인을 설계·구현하는 것을 목표로 운영된 세션이다. 데이터 생성부터 저장, 수집, 처리, 그리고 ETL 파이프라인 구축까지 단계적으로 학습하였으며, 이론 학습과 실습을 병행하는 방식으로 진행되었다.</p> <p>2) 주차별 활동 내용</p> <p>학기 초에는 Data Engineering 오리엔테이션을 통해 트랙의 전체 구성과 데이터 엔지니어의</p>		

역할에 대해 소개하였다. 이후 데이터 시스템과 데이터베이스 기초를 다루며 데이터 중심 애플리케이션 설계 개념과 데이터 저장 구조에 대한 이해를 쌓았다.

중반부에는 데이터 생성과 저장, 수집 및 처리 과정을 중심으로 학습을 진행하였다. 다양한 데이터 소스에서 데이터가 생성되는 방식과 이를 안정적으로 저장하고 관리하는 방법을 학습하였으며, 데이터 파이프라인의 핵심 구성 요소와 설계 원칙을 다루었다. 이 과정에서 데이터 품질과 신뢰성을 확보하기 위한 엔지니어링 관점의 접근을 함께 논의하였다.

중간고사 기간 이후에는 실습 중심의 수업이 진행되었다. ETL 개념을 기반으로 간단한 파이프라인을 직접 구현하며 데이터 추출, 변환, 적재 과정을 경험하였다. 실습은 Colab 환경과 로컬 환경에서 각각 진행되어, 개발 환경에 따른 차이와 운영상의 고려 사항을 비교할 수 있도록 구성되었다.

후반부에는 Airflow를 활용한 ETL 실습을 통해 스케줄링과 워크플로 관리 개념을 학습하였다. 또한 머신러닝 시스템 설계(MLOps) 관점에서 데이터 파이프라인이 모델 학습 및 운영과 어떻게 연결되는지에 대해 이해하였다. 학기 말에는 프로젝트를 완성하고, 전체 학습 내용을 정리하는 시간을 가졌다.

3) 활동 성과

본 트랙을 통해 참여자들은 데이터 엔지니어링의 전반적인 구조와 실제 데이터 처리 흐름을 체계적으로 이해할 수 있었다. 단순한 이론 학습을 넘어 직접 ETL 파이프라인을 구현해 봄으로써, 데이터 처리 과정에서 발생할 수 있는 문제와 이를 해결하는 방법에 대한 실무적인

감각을 기를 수 있었다.

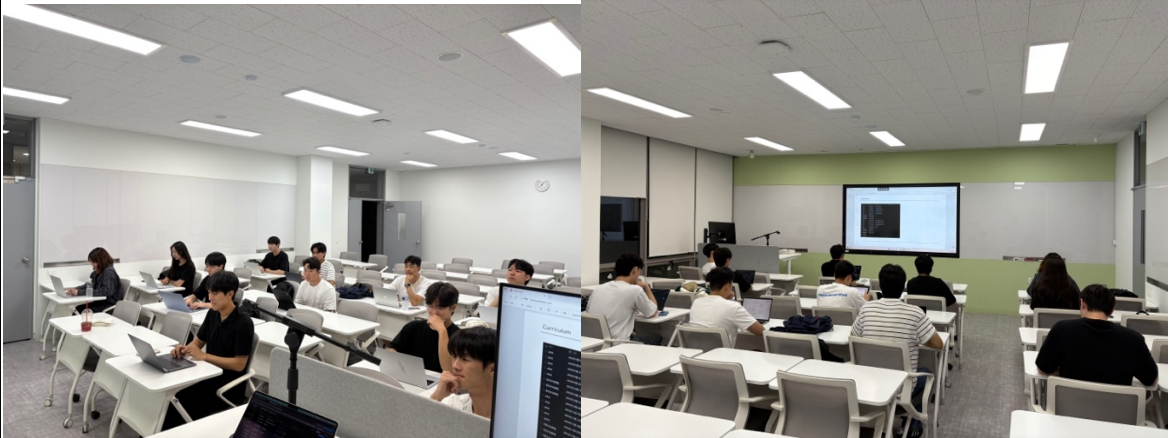
소감)

김XX 데이터가 생성되어 저장되고 처리되는 전 과정을 순서대로 학습할 수 있어 데이터 엔지니어링의 전체 그림을 이해하는 데 도움이 되었다. 실습 위주의 구성이라 이해도가 높았다.

이XX ETL 실습과 Airflow 활용을 통해 데이터 파이프라인이 실제로 어떻게 운영되는지 경험할 수 있었다. 이론과 실습이 잘 연결된 트랙이었다.

박XX 머신러닝 프로젝트에서 데이터 엔지니어링이 얼마나 중요한 역할을 하는지 체감할 수 있었다. 향후 심화된 데이터 처리 구조도 더 학습해 보고 싶다.

활동사진
(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬 박윤찬

동아리 활동보고서

활동명	NLP 심화트랙		
동아리 분과	학술	동아리명	KHUDA
활동시간	2025. 09. 01 ~ 2026. 01. 07	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 : 매 주차별 상이</p> <p>1) 활동 개요</p> <p>NLP 트랙은 한 학기 동안 자연어처리 및 대규모 언어모델(LLM)의 구조와 학습 과정을 단계적으로 학습하는 심화 세션으로 운영되었다. 이론 학습과 PyTorch 기반 구현, 미니 프로젝트를 병행하여 모델의 내부 구조와 실제 활용 방식을 이해하는 것을 목표로 하였다.</p> <p>2) 주차별 활동 내용</p> <p>학기 초에는 오리엔테이션과 함께 딥러닝 기초를 다루며 신경망 구조, 활성화 함수, 손실 함수, 옵티마이저 등 전반적인 딥러닝 기본 개념을 정리하였다. 이후 트랜스포머 구조를 중심으로</p>		

Self-Attention, Multi-Head Attention, Encoder-Decoder 구조를 학습하고, PyTorch의 nn.Transformer를 활용한 실습을 진행하였다.

중반부에는 아키텍처 구성 요소에 대한 심화 학습이 이루어졌다. Positional Embedding의 다양한 방식(Absolute, Relative, RoPE, ALiBi)을 비교하고, 데이터 정제 및 토큰라이저(BPE, WordPiece, SentencePiece) 차이에 따른 학습 특성을 분석하였다. 또한 Pretraining, Finetuning, RLHF, In-Context Learning 등 언어모델 학습 구조 전반과 Prompting과 Fine-tuning의 차이를 다루었다.

이와 함께 Efficient Training 기법을 학습하여 대규모 모델 학습 시 고려해야 할 자원 효율성과 학습 전략을 이해하였다. 이후 BLEU, ROUGE, Perplexity, HumanEval, LLM-as-a-judge 등 자연어처리 모델 평가 지표를 학습하고, 평가 방식의 한계와 활용 방안을 논의하였다.

후반부에는 국립국어원 벤치마크 데이터를 활용한 미니 프로젝트를 진행하였다. 토이 프로젝트 단계에서는 데이터 분석과 간단한 모델 적용을 수행하였으며, 중간 점검과 결과 공유를 통해 프로젝트 완성도를 높였다. 학기 말에는 심화 프로젝트를 수행하여 연구 아이디어 제안서 작성 및 컨퍼런스 형식의 발표를 진행하였다.

3) 활동 성과

본 트랙을 통해 참여자들은 자연어처리 모델의 구조적 이해부터 학습, 평가, 응용까지 전반적인 흐름을 체계적으로 학습할 수 있었다. 단순 이론 학습에 그치지 않고 구현과 프로젝트를 병행함으로써 실제 모델 동작 원리에 대한 이해도를 높일 수 있었다.

소감)

김XX 트랜스포머 구조를 처음부터 차근차근 구현하고 이해할 수 있어 자연어처리 모델에 대한 전반적인 구조가 정리되었다. 이론과 실습이 균형 있게 구성된 점이 인상 깊었다.

이XX 토큰나이저와 평가 지표 등 기존에 깊게 생각하지 않았던 요소들을 자세히 다룰 수 있어 도움이 되었다. 미니 프로젝트를 통해 배운 내용을 직접 적용해 볼 수 있었다.

박XX 한 학기 동안 LLM 학습 구조와 활용 흐름을 전반적으로 경험할 수 있는 트랙이었다.

심화 프로젝트 발표를 통해 연구 주제를 정리하는 연습이 된 점이 의미 있었다.

활동사진
(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬

박윤찬

동아리 활동보고서

활동명	최종 컨퍼런스		
동아리 분과	학술	동아리명	KHUDA
활동시간	2026. 01. 07	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 : 전자정보대학 211-1호</p> <p>1) 활동 개요</p> <p>심화프로젝트 컨퍼런스는 방학 및 학기 중 진행된 학습 활동의 최종 단계로, 팀별로 하나의 주제를 선정하여 장기간 심화 프로젝트를 수행하고 그 결과를 공유하는 형태로 진행되었다. 본 활동은 GitHub를 기반으로 한 협업과 프로젝트 관리 방식으로 운영되었으며, 각 팀은 데이터 분석, 머신러닝, 컴퓨터 비전, 자연어처리, 데이터 엔지니어링 등 서로 다른 기술 분야를 중심으로 프로젝트를 수행하였다.</p> <p>2) 팀별 프로젝트 주제 및 내용</p> <p>CV-DisabledParkingGuard 팀은 컴퓨터 비전 기술을 활용하여 장애인 주차구역을 보호하기 위한 시스템을 개발하였다. 이미지 또는 영상 데이터를 기반으로 주차구역과 차량을 인식하고, 장애인 주차구역 내 불법 주차 여부를 탐지하는 모델을 설계하였다. 데이터 전처리, 객체 탐지</p>		

모델 학습, 결과 시각화 과정을 포함하여 실제 적용 가능성을 고려한 프로젝트를 수행하였다.

NLP-CoverLetter 팀은 자연어처리 기법을 활용한 자기소개서 분석 및 문장 생성 프로젝트를 진행하였다. 입력된 자기소개서 텍스트를 분석하여 문장의 구조와 특징을 파악하고, 이를 바탕으로 보완 문장 또는 개선 방향을 제시하는 모델을 구현하였다. 텍스트 전처리, 임베딩, 언어모델 활용 등 자연어처리 전반의 과정을 프로젝트에 적용하였다.

NLP-Persona 팀은 사용자 페르소나에 따른 텍스트 생성 및 분석을 주제로 프로젝트를 수행하였다. 특정 성향이나 상황을 가정한 페르소나 정보를 기반으로, 이에 적합한 문장을 생성하거나 텍스트 특성을 분석하는 모델을 설계하였다. 다양한 페르소나 조건에 따른 결과 차이를 비교하며 모델의 활용 가능성을 검토하였다.

DB-Refocus 팀은 데이터베이스 구조 및 데이터 처리 방식 개선을 주제로 프로젝트를 진행하였다. 기존 데이터 구조를 분석하고, 쿼리 효율 및 데이터 접근성을 향상시키기 위한 개선 방안을 설계하였다. 데이터 관리 및 분석 효율을 높이기 위한 구조적 접근을 중심으로 프로젝트를 수행하였다.

MLFull-RIDER 팀은 머신러닝 전체 파이프라인을 적용하는 것을 목표로 프로젝트를 수행하였다. 데이터 수집 및 전처리부터 특징 추출, 모델 학습, 성능 평가 및 결과 해석까지 전 과정을 포함하여 머신러닝 프로젝트의 전반적인 흐름을 구현하였다. 모델 성능 비교와 해석을 통해 데이터 기반 인사이트 도출에 집중하였다.

DE-SteamQueue 팀은 데이터 엔지니어링 관점에서 대규모 데이터 처리 및 큐 기반 시스템을

주제로 프로젝트를 수행하였다. 데이터 흐름과 처리 병목을 분석하고, 효율적인 데이터 처리 파이프라인을 설계하여 시스템 구조 개선 방안을 제시하였다.

3) 활동 성과

각 팀은 프로젝트 수행 과정을 GitHub를 통해 체계적으로 관리하였으며, 코드와 문서를 정리하여 프로젝트 결과를 공유하였다. 이 과정에서 팀원 간 협업 경험을 쌓고, 각 기술 분야에서의 심화 학습을 진행할 수 있었다. 프로젝트 결과는 컨퍼런스 형식의 발표를 통해 공유되었으며, 질의응답과 피드백을 통해 프로젝트 완성도를 높였다.

소감)

이XX 심화 프로젝트를 통해 하나의 주제를 장기간 다뤄볼 수 있었고, 프로젝트 전반을 스스로 설계하고 구현하는 경험이 인상 깊었다.

최XX GitHub 기반 협업과 역할 분담을 통해 실제 프로젝트와 유사한 환경을 경험할 수 있었다. 기술적인 이해뿐만 아니라 협업의 중요성을 느낄 수 있었다.

박XX 기초부터 심화까지 이어진 활동의 마무리 단계로서 의미 있는 프로젝트였다. 각 팀의 발표를 통해 다양한 접근 방식을 접할 수 있었다.

활동사진
(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬

박윤찬

동아리 활동보고서

활동명	토이 프로젝트 발표		
동아리 분과	학술 분과	동아리명	KHUDA
활동시간	2025. 08. 27	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 : 전자정보대학 211-1호</p> <p>1) 활동 개요</p> <p>본 토이 프로젝트는 2025년 08월 27일에 진행된 실전 중심 활동으로, 앞서 진행된 머신러닝 기초 세션에서 학습한 개념을 실제 데이터에 적용해 보는 것을 목표로 하였다. 약 2주간의 준비 기간 동안 팀별로 주제를 선정하고, 데이터 분석부터 모델링, 결과 해석까지 전 과정을 직접 수행하였다.</p> <p>2) 활동 내용</p> <p>각 팀은 공개 데이터 또는 자체적으로 선정한 데이터를 활용하여 프로젝트를 진행하였다. 주요 활동 내용은 다음과 같다.</p> <p>문제 정의 및 분석 목표 설정</p> <p>데이터 수집 및 전처리 과정 수행</p> <p>기초 세션에서 학습한 머신러닝 모델 적용</p> <p>모델 성능 비교 및 결과 해석</p>		

분석 결과를 바탕으로 인사이트 도출

프로젝트 결과는 발표 자료로 정리하여 공유하였으며, 모델 선택 이유와 분석 과정, 결과 해석에 대해 팀별 발표와 질의응답을 진행하였다.

3) 소감

최 XX

기초 세션에서 배운 내용을 실제 데이터에 적용해 보면서 머신러닝 흐름을 전체적으로 이해할 수 있었다. 모델 성능뿐만 아니라 결과 해석이 중요하다는 점을 체감할 수 있는 프로젝트였다.

박 XX

이론으로만 접했던 개념을 직접 코드로 구현하고 결과를 비교해 보는 과정이 인상 깊었다. 팀원들과 함께 문제를 정의하고 해결해 나가는 경험이 의미 있었다.

박 XX

프로젝트를 통해 데이터 전처리의 중요성과 분석 과정의 어려움을 알게 되었다. 짧은 기간이었지만 실전 감각을 기를 수 있는 좋은 기회였다.

활동사진
(3매 이상)



동아리활동과 관련하여 상기 결과보고서를 확인합니다.

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

대표학생 : 박윤찬 박윤찬