

2024 KAIST XAI Tutorial Series

2024.11.05-11.22

최근 신뢰성 있는 AI 활용을 위한 규제와 입법이 본격화됨에 따라 AI 모델의 설명가능성 기술에 대한 관심이 그 어느 때보다 높아지고 있습니다. 이에 국내 최대의 설명가능 인공지능(XAI) 연구자 그룹인 KAIST 설명가능 인공지능연구센터(공동주최 성남연구센터)에서 2024 XAI 튜토리얼 시리즈를 개최합니다. 올해 프로그램은 XAI의 주요 알고리즘부터 최신 연구 주제까지 총 7일간 16개 발표세션으로 구성하였습니다.

많은 관심과 참여 부탁드립니다!



사전 참가신청

행사 홈페이지

<https://xai.kaist.ac.kr/Tutorial/2024>

참가대상

AI 관련 연구자, 학생, 산업계 종사자 및
AI에 관심있는 누구나

장 소 Hybrid (Zoom & KAIST 성남연구센터)
성남시 성남대로331번길 8, 킨스타워18층

주 최 KAIST 설명가능 인공지능연구센터
KAIST 김재철AI대학원 성남연구센터

DATE	TIME	SESSION	SPEAKER
Day 1 11.05(화)	15:00-15:50 16:00-16:50	1. 설명가능 인공지능 연구 동향 2. XAI Methods I : 지역적 설명 방법	교수 최재식 (KAIST/INEEJI) 연구원 이명진 (KAIST)
Day 2 11.07(목)	15:00-15:50 16:00-16:50	3. XAI Methods II : 전역적 설명 방법 4. XAI 평가지표	연구원 정영주 (KAIST) 연구원 김솔아 (KAIST)
Day 3 11.12(화)	15:00-15:50 16:00-16:50	5. 판별 언어모델의 설명: 주요 기법과 평가방안 6. 생성 언어모델의 설명: 프롬프팅 기반 설명	연구교수 김나리 (KAIST) 연구원 강청웅 (KAIST)
Day 4 11.14(목)	15:00-15:50 16:00-16:50	7. 생성모델과 XAI 8. 확산모델 기반의 이미지 생성AI에 대한 설명	연구원 최준호 (KAIST) 연구원 권다희 (KAIST)
Day 5 11.19(화)	15:00-15:50 16:00-16:50	9. 시계열 도메인 특화 XAI 기법 적용과 해석 10. XAI를 위한 인과관계	연구원 이세현 (KAIST) 연구원 조 원 (KAIST)
Day 6 11.21(목)	15:00-15:50 16:00-16:25 16:25-16:50	11. 메디컬 도메인에서의 XAI 12. 주요 XAI Framework 소개 13. Plug-and-Play XAI Framework 사용 튜토리얼	연구원 성지현 (KAIST) 연구원 이찬우 (KAIST) 연구원 김건형 (KAIST)
Day 7 11.22(금)	16:00-16:50 17:00-17:25 17:25-17:50	14. AI 규제 개요와 최신 동향 15. Understanding and Monitoring Model Behavior With Concept-based Explanations 16. Concept-based Explanations for LLM	변호사 이강혜 (법무법인 태평양) Maximilian Dreyer (Fraunhofer HHI) Reduan Achibat (Fraunhofer HHI)

문의 김나리 (nari.kim@kaist.ac.kr | 031-8022-7529)