

2022년 제1회 “K-ium” 의료 인공지능 경진대회 공고

1 추진배경

○ 목적

- 의료데이터 중심병원 지원사업을 통해 구축한 실전 데이터를 전국 대학생들을 대상으로 AI경진대회를 개최하여 인공지능 SW에 활용할 수 있는 기회를 제공하고자 함
- 의료데이터 중심병원 지원사업 데이터 관련 SW 개발을 지원하여 익명화 의료데이터 셋 공유 서비스 플랫폼 활성화에 기여
- 의료데이터 중심병원 지원사업 데이터를 활용한 우수 SW 개발을 통한 사회적 가치 창출

2 대회 개요

○ 대회명 : 2022년 제 1회 “K-ium” 의료 인공지능 경진대회

- 주관 : 부산대학교병원
- 주최 : 부산대학교병원, 전남대학교병원, 경북대학교병원
- 지원 : 보건복지부, 한국보건 의료정보원(의료데이터중심병원 지원 사업)

○ 공모부문 : 의료데이터 중심병원 지원 사업으로 구축된 익명화 의료 데이터 셋을 이용한 인공지능(AI) SW 개발

대회분야	익명화 의료 데이터 셋 활용 인공지능(AI) SW 개발
주요 내용	· 익명화 의료 Dataset을 활용한 인공지능 SW 개발 - CSV(Comma-separated values) 형식으로 공개 - 급성 허혈성 뇌졸중 판독문 (n=8843)

	- 한글 및 영어 혼용 자연언어 - 카테고리형 분류(급성 허혈성 뇌졸중 여부) - "Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International" (CC BY-NC-SA)의 라이선스에 따라 공개 - 크게 3개의 변수로 이루어져 있고 1번 변수는 영상소견을 적어둔 글자들(string), 2번 변수는 요약 소견의 글자들(string), 3번 변수는 이진변수로 급성/아급성 뇌경색이 있는 경우 1, 없는 경우 0으로 코딩 - 1차접수에서 datasets 70%만 공개(랜덤), 30% 2차 평가에서 검증
활용 데이터	· 익명화 의료데이터 셋 공유 플랫폼 활용(필수) - 익명화 의료 데이터 셋 공유 서비스 플랫폼(https://www.k-ium.com) DataSets

○ 참가자격 : 증명서를 제출할 수 있는 대학(원)생 누구나 참가 가능
(개인 또는 6인 이하로 구성된 팀, 중복참여 불가)

○ 추진일정 : '22년 9월~11월 (*코로나 19로 인해 일정 변동될 수 있음)

단 계	주요 일정	일정(안)
참가 접수 (1차 접수)	▪ 참가 접수: 2022.09.19.(월) ~ 2022.10.30.(일) 18시 마감 ▪ 회원가입 후 (https://www.k-ium.com) 참가신청 및 인공지능 모델 개발용 데이터소스 다운로드 ▪ 프로그램 결과물 압축파일 업로드 ※10.30.(일) 18시 마감	9월 19일(월)~ 10월 30일(일)
↓		
1차 평가	▪ 1차 선발: 2022.10.31.(월) ~ 2022.11.04.(금) ▪ 프로그램 결과물 1차 순위 선정	10월 31일(월)~ 11월 04일(금)
↓		
2차 접수	▪ 2차 접수: 2022.11.05.(토) ~ 2022.11.13.(일) 18시 마감 ▪ 검증데이터 공개, 결과 레포트 업로드 ※11.13.(일) 18시 마감	11월 05일(토)~ 11월 13일(일)
↓		
2차 평가	▪ 2차 선발: 2022.11.14.(월) ~ 2022.11.18.(금) ▪ 정확도(C-statistics AUROC) 검증, 레포트 평가	11월 14일(월)~ 11월 18일(금)
↓		
최종 발표	▪ 1차, 2차를 토대로 개발된 모델의 최우수 3개 팀(인) 선정 ▪ 온라인 발표 및 수상자 개별 통보	11월 23일(수)
↓		
시상식	▪ 진행방법: 비대면으로 진행, 수상자 대상 추후 별도 안내 ▪ 수상내용: 대학병원장상 및 부상품 수여, 언론보도 예정	11월 30일(수)

※ 일부 일정 및 내용은 변경될 수 있으며, 익명화 의료데이터 셋 공유 플랫폼에 공지함

3 접수

- 공모 및 접수 방법 : 익명화 의료데이터 셋 공유 플랫폼 (<https://www.k-ium.com>) 통한 공모 및 접수 (회원가입 필수)
- 접수 유의사항 : 참가 신청 시, 개인정보 수집 및 이용에 대한 동의와 데이터 셋 이용에 대한 동의 여부 필수 작성
(※붙임1.이용 동의서 및 붙임2.개인정보 수집 및 이용 동의서 참고)
- 제출 자료 : 프로그램 결과물 및 레포트 (기술 문서)
 - 1차 제출 : 프로그램 결과물(개발된 모델) 압축파일로 업로드
 - 2차 제출 : 결과물 검증을 위한 레포트 PDF형식으로 업로드

4 평가

- 평가방법 : 서류평가

구 분	내 용
평가 방법	▪ 1차 : 서류 심사(프로그램 결과물) ▪ 2차 : 서류 심사(레포트)
평가 위원	▪ 내·외부 관련 전문가 선정 (총 5인)
제출 규정 및 방법	(1) 공통 1차 제출 : 참가자는 접수 신청 이후 참가 접수 종료 시점 이전까지, 제공하는 가이드에 따라 프로그램 결과물을 업로드 하여야 함 2차 제출 : 참가자는 검증데이터 공개 이후 2차 접수 종료 시점 이전까지, 제공하는 가이드에 따라 레포트를 업로드 하여야 함 - 제출 결과는 파일을 수신하는 서버 시간상으로 제출 기한으로 규정된 시점까지 수신이 완료된 결과만이 인정되며, 이후에 수신된 결과는 인정되지 않음. ※ 최종 제출 성공한 파일만이 평가되므로

	올려야 할 파일을 잘못 선택하는 등의 업로드 실수가 없도록 필히 사전에 신중하게 확인 후 업로드 요망 (※참가자 실수로 인한 오제출에 대한 정정 신청은 받지 않음.) (2) 제출 규정 참가자는 제출양식을 따라 결과 파일을 생성하여 제출해야 하며, 양식 미준수로 인하여 발생할 수 있는 채점 오류나 제출 횟수 낭비 등의 불이익에 대한 책임은 참가자에게 있음. 결과는 반드시 주최 측에서 제공하는 서버 및 개발환경 내에서 지정한 방식을 이용하여 제출해야 하며, 외부에서 제출한 결과는 대회 참가 결과로 인정되지 않음.
평가 규칙	■ 정량평가 2차 평가에서는 정량적인 평가를 위해 C-statistics (AUROC)을 기반으로 정확도를 검증, 이를 기반으로 상위권 팀을 선정 ■ 정성평가 아래 내용을 포함한 개발 문서를 레포트 형태로 제출하게 되며, 이를 기반으로 서면 평가 진행. 동점자 처리를 위한 평가 항목이며 최대한 간략하게 A4 1장~10장 사이의 분량으로 작성하여 PDF 형식으로 제출하는 것을 권고 [정성평가 포함 내용] [제목] 제목, 팀명, 팀원 전체 목록 [요약] 적용한 기술의 전체적 요약 [방법] 적용한 기술의 자세한 방법 기술 [결과] AUROC 이외 추가적으로 제시할 수 있는 결과를 자유로운 형식으로 제시

○ 평가 기준

구분	검증 지표	점수
C-statistics(AUROC) 2차 평가를 기준	0-1 사이의 값으로 나타나게 되며, 소수 이하 6자리까지 비교 (소수 이하 7자리에서 반올림). 동점자 처리에 대해서는 상술 항목 참조.	100
합계		100

○ 대회 규정

1. 동점자 규정

2차 평가 시 C-statistics를 소수 이하 6자리까지 비교하여 높은 순으로 최종 시상. 1차 평가는 최종 점수에 반영하지 않음. 다만, 2차 평가에서 동점자가 발생할 경우 1차 평가 점수를 반영. 1차 평가의 점수를 계산할 때는 1차 데이터 전체를 개발된 모형에 적용하여 계산함. 1차 평가 점수까지 동점인 경우에는 정성 평가 항목으로 순위를 결정.

2. 윤리 규정

검증 시 부정행위(표절, 참가자로 기재되지 않은 사람의 조력 등 통상 학술 저술 활동에서 비윤리적이라고 판단하는 활동을 한 경우)가 의심되는 경우는 해당 팀 전원 탈락 처리함. 이는 주최 기관의 재량에 따라 부정 검증 프로그램, 제출된 코드 및 주식에 대한 개별적 심사 등의 방법으로 진행함. 인터넷 상에 공개된 코드나 라이브러리를 활용한 경우 출처를 소스 코드 내에 주식의 형태로 기재해야 함.

3. 공개 라이브러리 규정

공개 라이브러리가 아닌 라이브러리를 사용할 경우 윤리적 기준과 무관하게 탈락 처리함. 공개 라이브러리로 분류 가능한 라이선스의 예 및 공개 라이브러리의 예시는 아래와 같으며 참가자는 라이선스에 부과된 사용 조건을 엄수해야함.

- 공개 라이선스의 예: BSD, GNU GLP, LGPL, MIT, Apache licenses (전체 목록은 opensource.org의 웹페이지에서 제시하는 것에 준해서 판단함).

- 공개 라이브러리의 예: scikit-learn (BSD License), ML.net (MIT License), TensorFlow (Apache License), PyTorch (BSD-style variant)

4. 적용 가능한 프로그래밍 언어

기본적으로 자유롭게 선택 가능하며 모듈 간 다른 종류의 언어를 혼합하는 것도 가능하나 최대한 가독성이 높은 코드를 작성해야 하며 가독성이 떨어지는 부분에서는 주석을 충분히 사용할 것을 권장함. 다만 정성 평가 시 심사의 용이성을 고려하여 가능하다면 아래 계열의 프로그래밍 언어 사용을 권장하나 동점자가 없을 경우 심사 시 반영 요소는 아님.

- 권장 프로그래밍 언어: Python, C/C++/C#/Java, LISP/Scheme, R

5. 모형 검증 시 시간 제한 조건

제출된 모형은 최적화(hyperparameter tuning 포함)가 완료된 모형이어야 하며, 참가자가 사용하는 시가 5백만원 이하의 컴퓨터/워크스테이션 또는 이에 준하는 성능의 서버에서 10분 이내에 2차 심사 데이터셋을 처리할 수 있어야 함. 이는 참가자가 제출한 모형에 내재된 논리적 오류(무한루프, 지나친 재귀호출 등)로 인해서 심사에 어려움이 생길 정도로 많은 시간이 소요되는 것을 막기 위한 규정임.

5 시상 및 지원

○ 시상 내역

구 분	익명화 의료 데이터 셋 활용 인공지능(AI) SW 개발
최우수(3)	부산대학교병원장상 시상 (300만원 상당 부상)
	전남대학교병원장상 시상 (300만원 상당 부상)
	경북대학교병원장상 시상 (300만원 상당 부상)

※ 타인의 저작권을 침해하거나 유사대회에서 중복 수상한 경우 입상을 취소함

6

기타사항

○ 소유권 및 지식재산권

- 출품 및 수상모형은 저작권은 응모자에게 귀속되며, 주최기관 및 주관기관은 작품의 사용권을 가짐
- 신청·접수 이후 공개된 아이디어는 법적으로 보호받을 수 없으며, 이를 보호 받기 위해서는 신청자가 공개 이전에 직접 지식재산권을 획득하여야 함

○ 유의사항

- 개발된 모형과 데이터 셋(소스, 코드 포함) 모두 공공 개방 예정
- 타 경진대회에서 동일 결과물로 기 수상작 혹은 모방 작품으로 판명될 경우 심사에서 제외되며, 대회 종료 이후에도 시상 취소 및 상금회수
- 제출 서류상의 기재 착오나 연락불능으로 인한 불이익이 발생 하지 않도록 정보의 정확한 기재
- 경진대회의 전체 유의사항을 준수하며, 이를 준수하지 않을 경우 불이익
- 자격 유의사항
 1. 대학생 : 고등교육법 제 2조 제1호에서 제6호에 해당하는 학교(대학, 산업대학, 교육대학, 전문대학, 방송대학, 통신대학, 방송통신대학, 사이버대학, 기술대학)에 2022. 11. 30. 기준 재학 중인 자.
 2. 대학원생 : 고등교육법 제 29조에 따른 대학원(고등교육법 제 30조에 따른 대학원 대학이 두는 대학원 포함)에 2022. 11. 30. 기준 재학 중인 자.
 - ※ 재학생 범위는 재학생, 휴학생, 졸업유예생으로 증명서 발급이 가능한 자로 한정
 - ※ 시상식 일자를 기준으로 재학 시에 수상 자격을 인정.

7

문의처

- 문 의 처 : 익명화 데이터 셋 공유 서비스 플랫폼(<https://www.k-ium.com>)
게시판

1. 부산대학교병원(이하 ‘제공자’)이 본 대회에서 제공하는 데이터는 대회목적 및 대회기간에 국한하여 사용하여야 합니다.
2. 아래 행위를 포함하여 목적에 위배되는 일체의 행위를 할 수 없으며, 이에 위배되는 행동을 한 경우 제공자는 즉시 이용을 정지시킬 수 있습니다.
 - a. 데이터를 대회 플랫폼 밖으로 반출하거나 이를 시도하는 행위 (단, 개발 목적으로 개발 서버 또는 개인용 컴퓨터 내에 대회 기간 동안 일시적으로 데이터를 저장하는 것은 허용하며 대회 종료 시 삭제해야함. 이를 지키지 않아서 발생하는 데이터 유출과 관련된 책임은 참가자에게 있음.)
 - b. 대회 플랫폼을 제3자에게 양도하여 데이터에 접근하게 하는 행위
 - c. 데이터로부터 개인을 식별하려는 행위 및 의도적으로 대상자와 접촉하는 행위
 - d. 모자이크 혹은 블러 처리된 부분을 식별하려는 행위 및 기술적인 문제로 가려지지 아니한 부분을 의도적으로 찾아내려는 행위
 - e. 데이터의 원본 확인이 제한되는 상황에서 중간 결과를 대회 플랫폼에서 반출하거나 이를 이용하여 원본 데이터를 재구성하려는 행위
 - f. 기타 관계 법령에 위배되는 행위
3. 향후 논문 출판 등 연구성과를 발표할 경우 제공자로부터 해당 데이터를 제공받았음을 사사하여야 하며, 공개되지 않은 원본 데이터를 식별할 우려가 있는 결과물을 제3자 혹은 대중에게 노출할 경우 제공자와 사전에 협의해야 합니다.
4. 이용자는 데이터 및 메타데이터의 정확성에 대한 어떠한 보증 및 책임을 제공자에게 물을 수 없습니다.
5. 목적에 맞지 않는 이용 혹은 유출로 발생할 수 있는 일체의 법적 문제 및 불이익에 대해 제공자는 이용자에게 민형사상 책임을 물을 수 있습니다.

『2022년 제1회 k-ium 의료 인공지능 경진대회』 참가와 관련하여 『개인정보보호법』 제15조, 제17조, 제22조에 근거, 아래와 같이 개인정보의 수집 및 이용 목적을 위하여 개인정보의 수집 및 이용하는데 동의를 받고자 합니다.

☐ 정보의 수집 · 이용 목적

- 『2022년 제1회 k-ium 의료 인공지능 경진대회』의 대회 안내 및 참가, 개인 식별, 시상 등의 대회 운영의 제반사항을 목적으로 합니다.

☐ 정보 제공 받는자 : 주최 및 주관기관

☐ 수집 및 이용정보의 항목

- 개인식별정보(성명, 주민등록번호*, 주소*, 이메일, 휴대전화, 소속 등)
※ 상금 수령 시에 한하여 원천세 징수용으로 수집

☐ 보유 · 이용기간

- 위 개인정보는 수집 · 이용에 관한 동의일로부터 『2022년 제1회 k-ium 의료 인공지능 경진대회』 종료 후 1년까지 위 이용목적을 위하여 보유 · 이용됩니다.

☐ 동의를 거부할 권리 및 동의를 거부할 경우의 불이익

- 지원자는 위와 같이 개인정보 중 필수항목의 수집 · 이용에 관한 동의를 거부할 권리가 있습니다. 동의를 거부할 경우 『2022년 제1회 k-ium 의료 인공지능 경진대회』 참가선발 시 불이익이 있을 수 있습니다.

☐ 정보의 파기

- 수집된 개인정보는 안전하게 처리하며, 이용 및 보유 기간이 지난 정보는 삭제 및 파기합니다.

상기 본인은 팀의 대표로써, 참가를 위해 위와 같이 개인정보 수집 및 이용에 동의함