

질병분류기호 추천 딥러닝 모델

KUNLP팀

건국대학교 컴퓨터공학부
정근영, 윤동근

질병분류기호에 대한 설명

- 한국 표준 질병사인분류(Korean Standard Classification of Diseases, KCD)
 - 질병이환 및 사망자료를 유사성에 따라 체계적으로 유형화
 - 모든 형태의 보건 및 인구동태 기록에 기재되어 있는 질병 및 기타 보건문제를 분류하는데 이용하기 위하여 설정한 것

☞ 병원 뿐만 아니라 약국, 보험 등
범용적으로 쓰이는 질병 코드

- ▶ [A00-B99] I. 특정 감염성 및 기생충성 질환(A00-B99)
- ▶ [C00-D48] II. 신생물(C00-D48)
- ▶ [D50-D89] III. 혈액 및 조혈기관의 질환과 면역매커니즘을 침범한 특정 장애(D50-D89)
- ▶ [E00-E90] IV. 내분비, 영양 및 대사 질환(E00-E90)
- ▶ [F00-F99] V. 정신 및 행동 장애(F00-F99)
- ▶ [G00-G99] VI. 신경계통의 질환(G00-G99)
- ▶ [H00-H59] VII. 눈 및 눈 부속기의 질환(H00-H59)
- ▶ [H60-H95] VIII. 귀 및 유도의 질환(H60-H95)
- ▶ [I00-I99] IX. 순환계통의 질환(I00-I99)
- ▶ [J00-J99] X. 호흡계통의 질환(J00-J99)
- ▶ [K00-K93] X I. 소화계통의 질환(K00-K93)
- ▶ [L00-L99] X II. 피부 및 피하조직의 질환(L00-L99)
- ▶ [M00-M99] X III. 근골격계통 및 결합조직의 질환(M00-M99)
- ▶ [N00-N99] X IV. 비뇨생식계통의 질환(N00-N99)
- ▶ [O00-O99] X V. 임신, 출산 및 산후기(O00-O99)
- ▶ [P00-P96] X VI. 출생전후기에 기원한 특정 병태(P00-P96)
- ▶ [Q00-Q99] X VII. 선천기형, 변형 및 염색체이상(Q00-Q99)
- ▶ [R00-R99] X VIII. 달리 분류되지 않은 증상, 징후와 임상 및 검사의 이상소견(R00-R99)
- ▶ [S00-T98] X IX. 손상, 중독 및 외인에 의한 특정 기타 결과(S00-T98)
- ▶ [U00-U99] X X II. 특수목적 코드(U00-U99)
- ▶ [V01-Y98] X X. 질병이환 및 사망의 외인(V01-Y98)
- ▶ [Z00-Z99] X X I. 건강상태 및 보건서비스 접촉에 영향을 주는 요인(Z00-Z99)

주제 선정 배경

- 건국대학교 제공 초진 진료 데이터
- 6,000개 이상의 진단 코드
- 건대병원에서는 진단 코드 입력하면
진단명, 진단분류까지 매칭
☞ 의사가 진료 과정에서 코드를 선택
- 진단코드가 진료 과정에서 가장 중요한 내용

주호소 및 현병력	진단코드	진단명	진단분류	진단소분류	진료과
GB polyp	K828-001	GB polyp	담낭의 기타 명시된 질환	담낭의 기타 질환	소화기내과
개인의원에서 시행한 검진 복부 초음파 검사에서 담낭 내 용종이 관찰되어 내원함	K828-001	GB polyp	담낭의 기타 명시된 질환	담낭의 기타 질환	소화기내과
AF	I639-001	Cerebral Infarction	상세불명의 뇌경색증	뇌경색(증)	심장혈관내과
10/31 AF 으로 의뢰 현재 요양병원 거주	I639-001	Cerebral Infarction	상세불명의 뇌경색증	뇌경색(증)	심장혈관내과
consult f/u	H2692	Cataract, unspecified, Bilateral	상세불명의 백내장, 양쪽	기타 백내장	안과
DM- HT- 2019년 4월부터 좌안 시력저하 ** 양안 산동 good	H2692	Cataract, unspecified, Bilateral	상세불명의 백내장, 양쪽	기타 백내장	안과

주제 선정 배경

- 의사의 진료 과정에서의 직접적인 도움
- 진료 과정 효율 증가

주호소 및 현병력	진단코드	진단명	진단분류	진단소분류	진료과
GB polyp	K828-001	GB polyp	담낭의 기타 명시된 질환	담낭의 기타 질환	소화기내과
개인의원에서 시행한 검진 복부 초음파 검사에서 담낭 내 용종이 관찰되어 내원함	K828-001	GB polyp	담낭의 기타 명시된 질환	담낭의 기타 질환	소화기내과
AF	I639-001	Cerebral Infarction	상세불명의 뇌경색증	뇌경색(증)	심장혈관내과
10/31 AF 으로 의뢰 현재 요양병원 거주	I639-001	Cerebral Infarction	상세불명의 뇌경색증	뇌경색(증)	심장혈관내과
consult f/u	H2692	Cataract, unspecified, Bilateral	상세불명의 백내장, 양쪽	기타 백내장	안과
DM- HT- 2019년 4월부터 좌안 시력저하 ** 양안 산동 good	H2692	Cataract, unspecified, Bilateral	상세불명의 백내장, 양쪽	기타 백내장	안과

주제 소개

질병분류기호 추천 딥러닝 모델

주호소 및 현병력 (의사 입력)

- 6개월 전 부비동염 수술 받음 개인병원. 미열, 두통, 코막힘 항상, 후비루 있어 내원함. 현재 항생제는 복용하지 않고 있음. 알레르기 검사 음성이었다고 함

진단 코드(모델 출력)

- (R0980) Nasal congestion - 코울혈
- (J310) Chronic rhinitis - 만성 비염
- (R070) Pain in throat - 목구멍의 통증
- (R093) Abnormal amount of sputum - 이상가래
- (J339) Nasal polyp, unspecified - 상세불명의 코폴립

주제 소개

질병분류기호 추천 딥러닝 모델

주호소 및 현병력 (의사 입력)

- 안면부 홍반 구진 면포. 세레스톤지를 1년 가까이 발랐다고 함. 중단하세요. 중단후 rebound 현상 나타날 수 있음을 설명함.

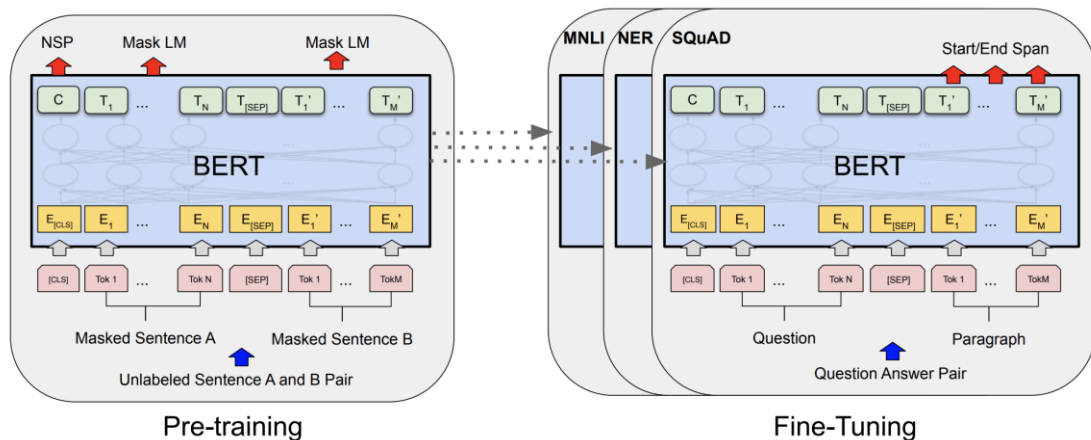
진단 코드(모델 출력)

- (L709) Acne - 상세불명의 여드름
- (L739) Folliculitis - 상세불명의 모낭장애
- (L719) Rosacea, unspecified - 상세불명의 주사
- (L219) Seborrheic dermatitis - 상세불명 지루피부염
- (L720) Epidermal cyst - 표피낭

기술적 배경

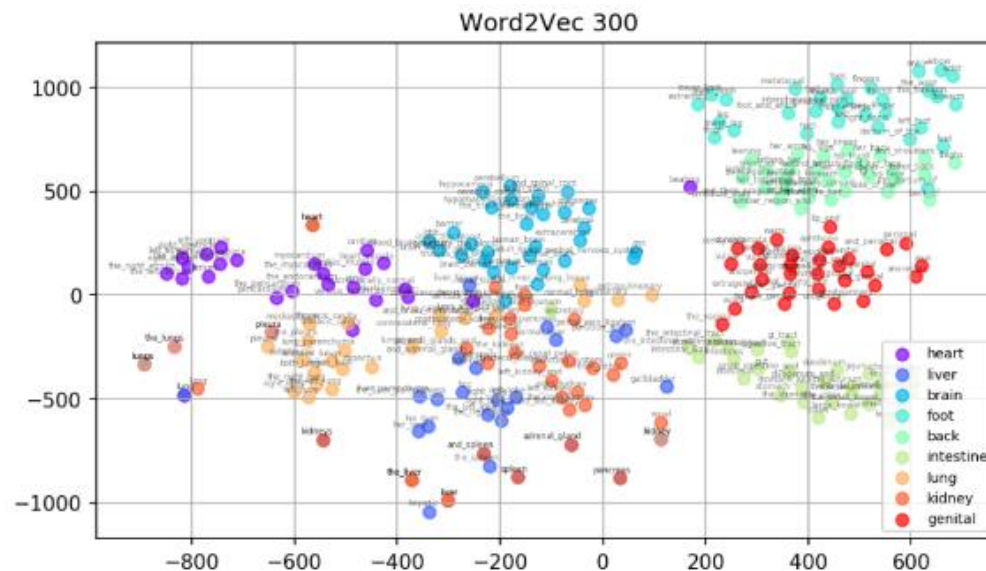
Roberta

- 자연어처리 AI 언어모델
(62GB의 한국어 데이터로 사전학습된 언어모델)

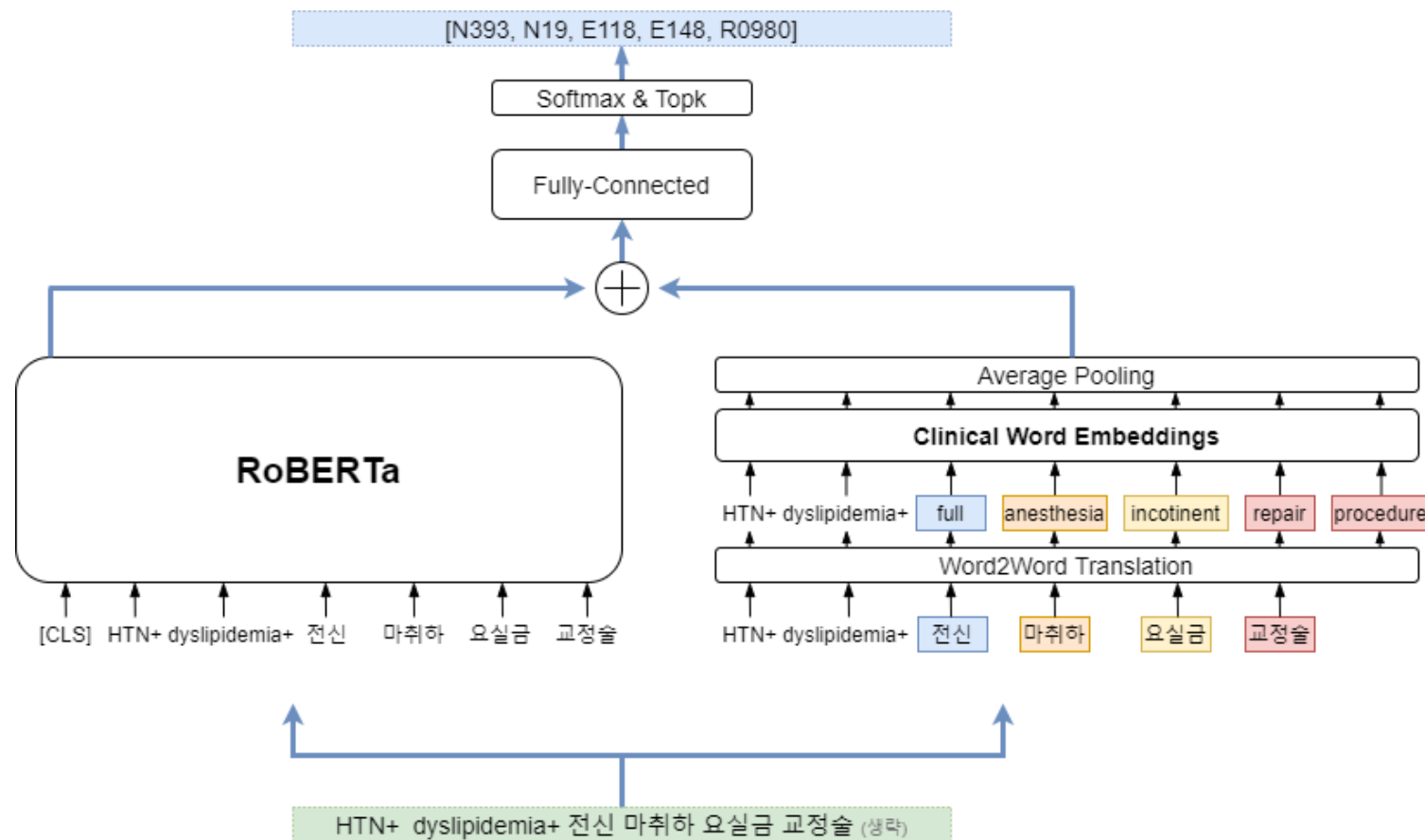


Clinical Word Embeddings

- PUBMED data(122GB의 의학 데이터)로 의학용어
특징 학습



모델 구조



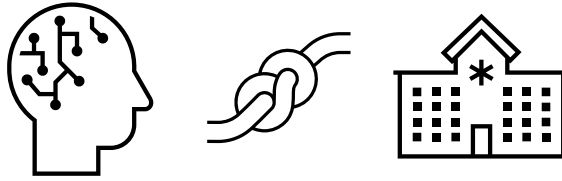
- 데이터 개수 : 13만개
- 95%의 데이터를 학습,
5%의 데이터를 평가 데이터로 사용
- GPU : RTX 8000
- Epochs : 15
- 학습시간 : 22시간 (1시간 30분 * 15)
- Model 파라미터 개수 : 3억 4천만

모델 성능

Top-k Accuracy : 입력된 증상에 대해 모델이 결과로 내놓은 k개의 질병 코드 중,
실제 정답 질병 코드가 포함되어 있을 확률(정확도)

	1	3	5
Top-k Accuracy	50%	72%	80%

창업 계획



- 병원의 데이터를 적용한 별도의 모델 필요
 - 병원의 진료 프로그램과 결합
- 특정 병원의 필요에 맞는 서비스 제공 가능
 - 모델 솔루션 제공

데모 시연

203.252.166.40:3000

질병분류코드 추천 딥러닝 모델

2021 KU 메디컬 해커톤 KUNLP 팀

건국대학교 컴퓨터공학부 정근영, 윤동근

환자의 주소록 및 현병력

6개월 전 부비동염 수술 받은 개인병원. 미열, 두통, 코막힘 증상, 후비루 있어 내원함. 현재 항생제는 복용하지 않고 있음. 알레르기 검사 음성이었다고 함

진단코드 추천

R0980

Nasal congestion

코골임

J310

Chronic rhinitis

만성 비염

R070

Pain in throat

목구멍의 통증

R093

Abnormal amount of sputum

이상가래

J339

Nasal polyp, unspecified

상세포종의 코폴립

Q&A