



ToBigs 7th Conference - TMI 팀

Q

통합검색

፧달 카테고리 보기 ▼

김동규 박규리 신훈철

장유영 장청아 최세영

도움말

자동완성 끄기

Executive Summary

본 연구의 목적은 질문에 대한 간결한 답을 얻어내는 것임

한국형 SQuAD 데이터와 지식백과 데이터를 기반으로 하여, 문서 유사도, NER, POS, Self-Attention 개념을 조합하여 한국어판 QA 모델을 제작함

본 모델의 경우 고차원적인 추론은 불가능하지만 일반적인 상식 수준의 질문에 대한 성능은 우수하였음

토큰화 및 개체 명 인식 등 세부 TASK들에서의 추가 연구가 필요하다고 판단됨

Sele	ect or Type Yo	our Options	
Chap	ter		
Ch1	. 주제 선정 배경		
Ch3 Ch4 Ch5	. 모델 Overview . 데이터 수집 및 정 . 모델 세부 구조 . 결과 및 QA봇 서 . 결론		
Date			
2 nd R	ound Place		
	Accept	Cancel	

장영실이 만든 물시계의 이름



통합검색 블로그 카페 지식iN 이미지 동영상 어학사전 □ 뉴스 더보기 ▼ 검색옵션 🗸

웹사이트

도움말

물시계 - 나무위키

https://namu.wiki/w/물시계 -

이름은 Rhei. 3. 관련 문서 [편집] 자격루 최근 변경 최근 토론 특수 기능 라이브 최근 변경 00:08오다카역(후 쿠시마현)... [더 보기] 0 역링크 토론 편집 역사 ACL물**시계** 최근 수정 시각: 2017-04-23 06:18:09 분류 시 계 이 저작물은 CC BY-NC-SA 2.0 KR에...

나무위키

이창규 기자 - '옥루'가 도대체 뭐길래?…장영실이 만든 천상(天象) 자동 물....

http://www.topstarnews.net/news/articleView.html?idxno=497788

옥루는 1438년(세종 20년)에 **장영실이 만**든 천상(天象) 자동 **물시계의 이름**이다. 1434년(세종 16년) 자격루 를 완성한 장영실은 천체의 운행을 관찰하는 혼천의와... …장영실이 만든 천상(天象) 자동 물시계 이창규 기 자 승인 2018.10.11 00:52 댓글...

톱스타뉴스

'옥루'가 도대체 뭐길래?…장영실이 만든 천상(天象) 자동 물시계

https://blog.naver.com/topstarnews/221375120475

옥루는 1438년(세종 20년)에 장영실이 만든 천상(天象) 자동 물시계의 이름이다. 1434년(세종 16년) 자격루 를 완성한 **장영실**은 천체의 운행을 관찰하는 혼천의와 시간을 알려주는 자격루를 결합한 천문기구를 만들고자 했다. 혼천의 / 뉴시스 김돈(金墩)이 쓴...

네이버 블로그 | 2018.10.11.



Q

통합검색 블로그 카페 지식iN 이미지 동영상 어학사전 □ 뉴스 더보기 ▼

검색옵션 🗸

웹사이트

도움밀

물시계 - 나무위키

https://namu.wiki/w/물시계 ▼ 이름은 Rhei. 3. 관련 문서 [편Q: ス장영실이 코다든 물시계의 이름은?변경 00:08오다카역(후 쿠시마현)... [더 보기] 0 역링크 토론 편집 역사 A/주시·격루수정 시각: 2017-04-23 06:18:09 분류 시계 이 저작물은 CC BY-NC-SA 2.0 KR에...

- idea-

이 장규 기자질문에 대한 답을 간결하게 출력하는 시스템을 만들고자 함

http://www.topstarnews.net/news/articleView.html?idxno=497788

옥루는 1438년(세종 20년)에 **장영실이 만**든 천상(天象) 자동 **물시계의 이름**이다. 1434년(세종 16년) 자격루를 완성한 **장영실**은 천체의 운행을 관찰하는 혼천의와... **…장영실이 만든** 천상(天象) 자동 **물시계** 이창규 기자 승인 2018.10.11 00:52 댓글...

톱스타뉴스

'옥루'가 도대체 뭐길래?…장영실이 만든 천상(天象) 자동 물시계

https://blog.naver.com/topstarnews/221375120475

옥루는 1438년(세종 20년)에 **장영실이 만**든 천상(天象) 자동 **물시계의 이름**이다. 1434년(세종 16년) 자격루를 완성한 **장영실**은 천체의 운행을 관찰하는 혼천의와 시간을 알려주는 자격루를 결합한 천문기구를 만들고자했다. 혼천의 / 뉴시스 김돈(金墩)이 쓴...

네이버 블로그 | 2018.10.11

사람의 독해 능력을 뛰어 넘고자 하는 현재 NLP의 트렌드와 편리한 검색에 대한 사용자의 니즈를 고려하여 QA 시스템을 도출하고자 함

 $- \times$

(1) 검색 편리성 증대

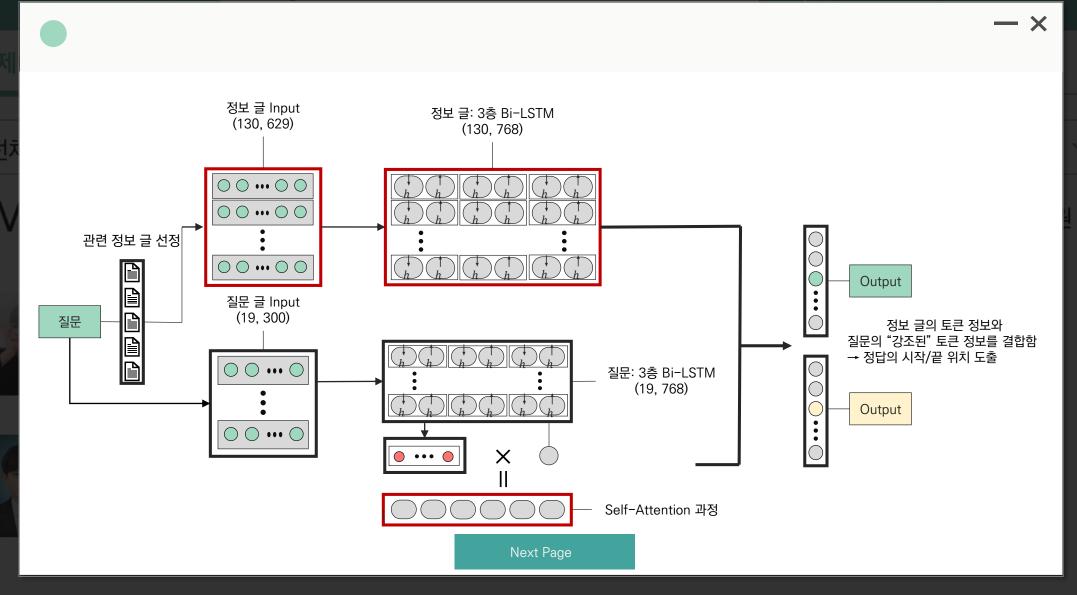
(2) 현 NLP 분야의 발전 방향

질문에 대한 답을 간결하게 얻고자 하는 사용자의 Unmet Needs가 있음 사람의 독해 능력을 능가하는 모델을 구현하는 일은 현재 NLP 분야에서 주목을 받고 있음

QA 시스템: TMI 서비스 구축

Chapter			
Ch1. 주제 선정			
Ch2. 모델 Ove Ch3. 데이터 수			
Ch4. 모델 세부			
Ch5. 결과 및 (=	: 시연	
Ch6. 결론			
Date			
2 nd Round Plac	ce		

본 연구에서 적용한 DrQA 모델의 전체적인 구조는 아래와 같음 ToBigs 7th Conference - TMI 팀



Chapte	' F제 선정 배경		
	면 Overview		
	베이터 수집 및 정지	41	
	2델 세부 구조 열과 및 QA봇 서비	Д Л О	
Ch6. ?	-	스 시킨	
Date ———			
2 nd Rou	ınd Place		



LG CNS Korquad

Context : 10,645개

QA 쌍 : 66,181개

AI HUB Kowiki

Context : 61,801개

QA 쌍 : 100,268개

네이버 지식백과

인물 및 요약: 31842개

Train/ Dev set

Test set



id

('8_C58_wiki_327-1',

context

'그러나 조조는 암살에 실패하여 진류로 달아나 각지의 제후들과 함께 반동탁 연합군을 일으켰다. 연합군에 밀려 장안으로 천도한 후에도 동탁의 행동이 달라지지 않자, 왕윤은 자신의 수양딸 초선으로 하여금 동탁과 여포의 사이를 이간질시키고 두 사람의 사이를 벌어지게 만들었다. 여포가 동탁을 죽일 결심을 굳히자, 이숙을 동탁에게 보내어 현제가 동탁에게 선양하려 한다고 전하고 궁궐로 불러들여 동탁을 죽였다. 이것을 연환지계라 한다. 그러나 동탁의 잔당이 장안을 공격해 오자 장안을 빼앗기고 이각에게 붙잡혀 죽었다. 왕굉(왕굉 (장문)) 왕릉(왕릉 (조위)) 왕윤(王允, 137년 ~ 192년 7월 4일(음력 6월 7일))은 중국 후한 말의 정치가로, 자는 자사(子師)이며 병주(병주 (중국))(幷州) 태원군(타이위안 시)(太原郡) 기현(祁縣) 사람이다. 여포(呂布)를 움직여 전횡을 일삼던 동탁(董卓)을 죽였으나, 동탁의 하인 이각과 곽사의 반격을 받고 헌제 앞에서 죽었다.',

question

· '동탁과 여포 사이를 이간질 한 사람은 누구지',

answer

'왕윤',

326,

328)

네이버 지식백과 Crawling 데이터

ToBigs 7th Conference - TMI 팀



Title	Context	Summary
프레디 머큐리	1946년 탄자니아 잔지바르에서 태어났다. 어린 시절 인도로 이주해 학교를 다녔고, 12세 때 스쿨 밴드 헥 틱스(The Hectics)를 결성해 클리프 리처드(Cliff Richard), 리틀 리처드(Little Richard)등 로큰롤 가 수들의 노래를 커버했다. 1964년 영국으로 이주했고, 1960년대 후반 아이벡스(Ibex), 사우워 밀크 시 (Sour Milk Sea) 등의 그룹에서 활동했다.	프레디 머큐리는 영국의 록 보컬리스트. 프레디 머큐리는 1946년 탄자니아 잔지바르에서 태어 남. 프레디 머큐리는 1960년대 후반 아이벡스 (lbex), 사우워 밀크 시(Sour Milk Sea) 등의 그룹에서 활동함. 프레디 머큐리는 1985년 첫 번째 솔로 앨범 《미스터 배드 가이》(1985년 영 국 앨범차트 6위), 1988년 스페인 출신의 소프 라노 몽세라 카바예(Montserrat Caballé)와 함 께 만든 앨범 《바르셀로나(Barcelona)》(1988 년 영국 앨범차트 15위)를 연이어 발표함.
윤동주	만주 북간도의 명동촌(明東村)에서 태어났으며, 기독교인인 할아버지의 영향을 받았다. 본관은 파평(坡平)이며, 아버지는 윤영석(尹永錫), 어머니는 김룡(金龍)이다. 1931년(14세)에 명동(明東)소학교를 졸업하고,한 때 중국인 관립학교인 대랍자(大拉子) 학교를 다니다 가족이 용정으로 이사하자 용정에 있는 은진(恩眞)중학교에 입학하였다(1933).1935년에 평양의 숭실(崇實)중학교로 전학하였으나, 학교에 신사참배 문제가 발생하여 폐쇄당하고 말았다.	윤동주는 기독교인인 할아버지의 영향을 받음. 윤동주는 1931년(14세)에 명동소학교를 졸업 함. 윤동주는 다시 용정에 있는 광명학원의 중학 부로 편입하여 거기서 졸업함. 윤동주는 1941 년에는 서울의 연희전문학교 문과를 졸업함. 윤 동주는 학업 도중 귀향하려던 시점에 항일운동 을 했다는 혐의로 일본 경찰에 체포됨. 윤동주는 그의 시집은 본인이 직접 발간하지 못함.



Tokenizing

Twitter/Komoran/Kkma 중 모델에 적합한 Twitter 사용

Regular Expression

불용어가 많아 분석에 있어 필수적인 '~', '.', 한글, 영어, 숫자 만을 추출함

Embedding

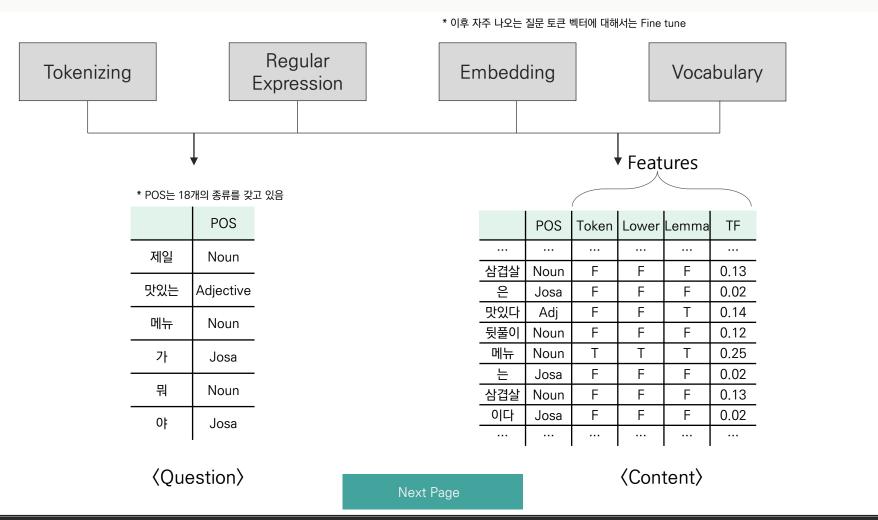
200차원, CBOW에 비해 300차원, Skipgram의 성능이 우수하였음

Vocabulary

	 Unk	Pad	배	가	고프	다	
_	 0	1	500	4056	331	6948	
		'				' * Index 는	' 자주 나오는 토큰부터 오름차순

데이터 전처리 결과

ToBigs 7th Conference - TMI 팀



Select or Ty	pe Your Options	
Chapter		
Ch1. 주제 선정	배경	
Ch2. 모델 Over Ch3. 데이터 수		
Ch4. 모델 세부	구조	
Ch5. 결과 및 Q Ch6. 결론	A봇 서비스 시연	
Date		
2 nd Round Place	е	
Accept	Cancel	

모델은 크게 3가지의 핵심 요소로 구성되어 있음: DR, NER, QA



사용자의 질문에 적합한 답을 찾기 위해서는 그와 관련된 정보 글 (Document)을 찾아내는 과정은, 문장-정보 글 사이의 유사도 계산을 통해 진행됨

질문, 정보 글 예시 **── ★**

질문

카보베르데는 누가 발견했는가?

정보 글

카보베르데 공화국(카보베르데共和國, ,), 줄여서 카보베르데는 아프리카 서쪽 대서양에 있는 국가이다. 영어로 케이프베르데라는 이름으로도 알려져 있다. 포르투갈의 항해자들에 의해 발견되었다. …

유사도 계산

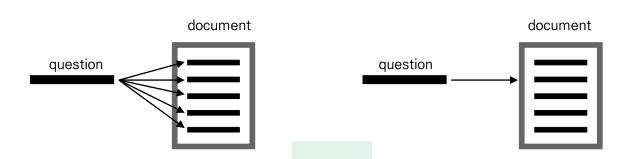
$$W_{x,y} = t f_{x,y} \times \log(\frac{N}{df_x})$$

 $tf_{x,y}$ = frequency of x in y df_x = number of documents containing xN = total number of documents

- TF-IDF: Term x within document y
- TF-IDF 기반의 코사인 유사도를 사용함
- 벡터 공간에서 각도를 보고 유사도를 판단하는 수치임
- 문서의 길이에 상관없이 정규화 가능

유사도 계산 세부 과정: 질문과 정보 글의 유사도를 계산한 방법은 아래와 같음





문장 단위	정보 글 전체 단위
TF-IDF – cosine similarity	TF-IDF – cosine similarity
Bigram-TF-IDF - cosine similarity	Bigram- TF-IDF – cosine similarity
TF-IDF x token embedding - cosine similarity	TF-IDF x token embedding - cosine similarity

질문과 정보 글(Document)의 유사도 핵심 키워드는 특별 고유명사임

Next Page

 $- \times$

유사도 계산 세부 과정: 정보 글을 찾을 때 특별 고유명사만 뽑아서 만든 유사도에 가중치를 더 주는 과정

 $- \times$

특별 고유명사가 포함된 글 예시

질문

1961년 투쟁은 나의 인생이라는 자서전을 낸 대통령이자 흑인인권운동가는?

정보 글

넬슨 만델라는 남아프리카공화국 최초의 흑인 대통령이자 흑인인권운동가임. 넬슨 만델라는 1918년 7월 18일 트란스케이 움타타에서 출생함. …넬슨 만델라는 저서로 자유를 위한 투쟁 의지를 밝힌 《투쟁은 나의 인생 The Struggle is MyLife》(1961)과 자서전 《자유를 향한 머나먼 여정 Long Walk to Freedom》(1995)이 있음

[특별 고유명사에 가중치를 부여하는 방법]

전체 명사의 TF-IDF로 구한 코사인 유사도 + 특별 고유 명사의 TF-IDF로 구한 코사인 유사도 * 2 위 식 계산 후 가장 높은 5개 추출

특별 고유명사 반영 X

특별 고유명사 반영 ○

질문:하늘과	והוצום	เลาเ	11015	ᄁ	그녀 중단중단	자니노
> T - U = F	마음판	> P	ALC: T	\wedge I $>$	710121	ニジノロー

담변

28878

27954

열자는 BC 400년경 정나라에 살았다고 전하나 《사기》에는 그 전기가 보이지 않... 보라이언 올디스는 영국 대표 양작가.《목스퍼드 메일》지의 문학란 담당 편집자,영... 캘러현는 아일랜드계 캐나다의 작가로서 사회주의와 그리스도교적 사랑에 대한 말도 높... 양먹는 중국 북송의 시인. 양먹는 문재가 뛰어나 11세에 천자로부터 시재를 인정받... 정사위는 자는 홍원이고, 하는 병은의, 정사위는 1564년(명종 19) 사마시에 ... 질문:하늘과 바람과 별과 시라는 시를 편찬한 작가는

_⊦₩

26891 윤동주는 만주 북간도의 명동촌에서 태어남. 윤동주는 일본으로 건너가 도쿄에 있는 ... 29532 카자케비치는 극동지방에서 문화지도위원 등에 종사하다가 1938년 모스크바로 돌아와... 29911 베르텔 그리펜베리는 20세기 초, 러시아 태생의 핀란드 시인. 베르텔 그리펜베리는... 26602 보리스 라브레뇨프는 러시아의 소설가 · 극작가. 보리스 라브레뇨프는 모스크바대학교 . 28381 클로드 시몽는 프랑스의 소설가로 누보 로망(반소설)의 대표적 작가임. 클로드 시몽...

개체 명 인식(NER: Named-entity recognition)의 필요성 ToBigs 7th Conference - TMI 팀



개체 명

문서에서 특정한 의미를 가지고 있는 단어 또는 어구, 정보 검색에서 중요 검색 대상이 됨

QA 시스템에서의 개체 명

사람, 지역, 날짜, 시간, 기관 등 개체 명은 context 내에서 주요 키워드일 가능성이 높음

context

한국인 선수인 손흥민은 토트넘 핫스퍼에서 축구를 한다.

〈나라〉

〈사람〉

〈기관〉

〈문명〉

한국인 선수인 기성용은 뉴캐슬 유나이티드에서 축구를 한다.

〈나라〉

〈사람〉

〈기관〉

〈문명〉

총 5개의 개체 명 인식 Task 진행

question

answer

토트넘 핫스퍼에서 축구 선수 생활을 했던 한국 사람은? 〈기관〉 〈문명〉

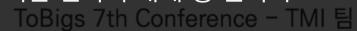
〈나라〉

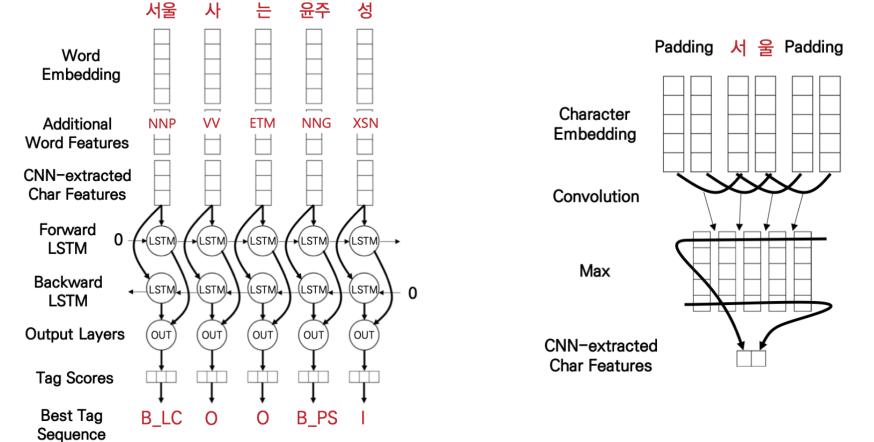
손흥민 〈사람〉

Next Page

개체 명(영문) 개체 명(국문) **PERSON** 사람 **FIELD** 학문 대상물 ARTIFACTS WORKS 기관 **ORGANIZATION** 지역 LOCATION 문명 CIVILIZATION 날짜 DATE TIME 시간 숫자 NUMBER **EVENT** 사건/사고 동물 **ANIMAL** 식물 **PLANT** 금속 **MATEPIAL** 용어 **TERM**

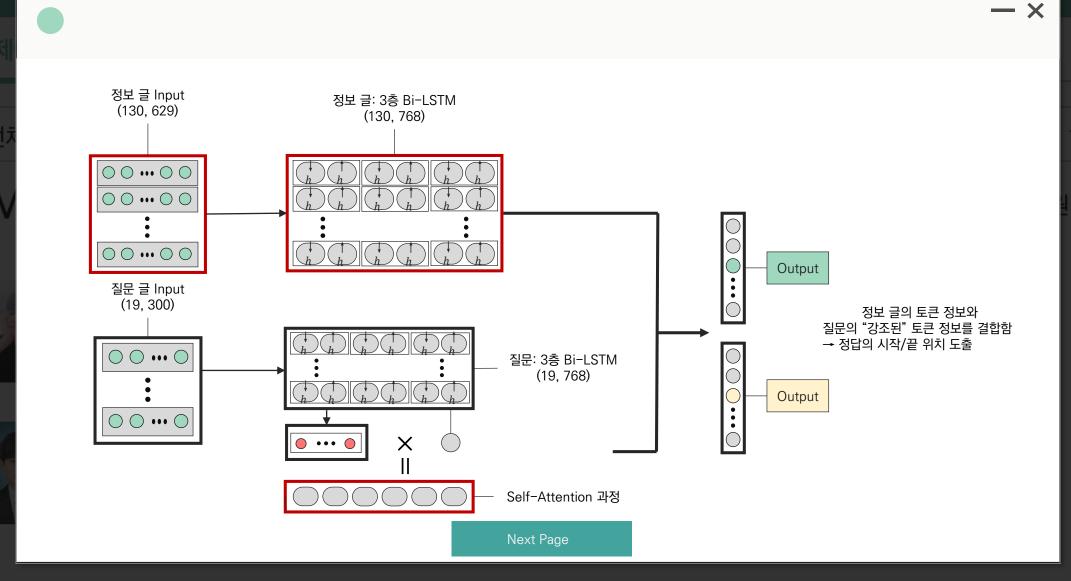
CNN + Bi-LSTM 기반 한국어 개체 명 인식기



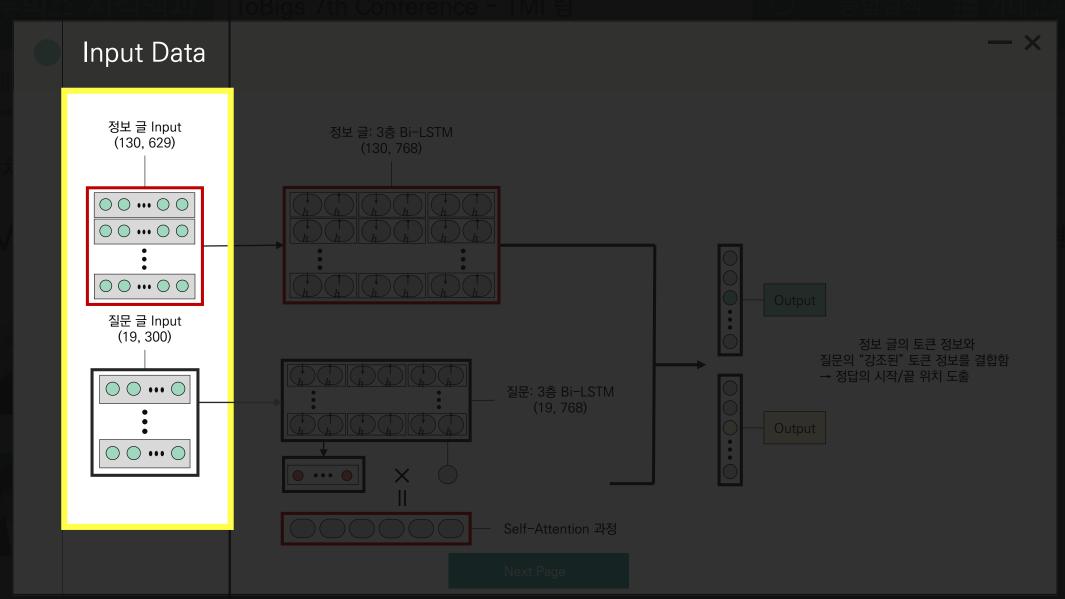


QA 모델 전체 구조

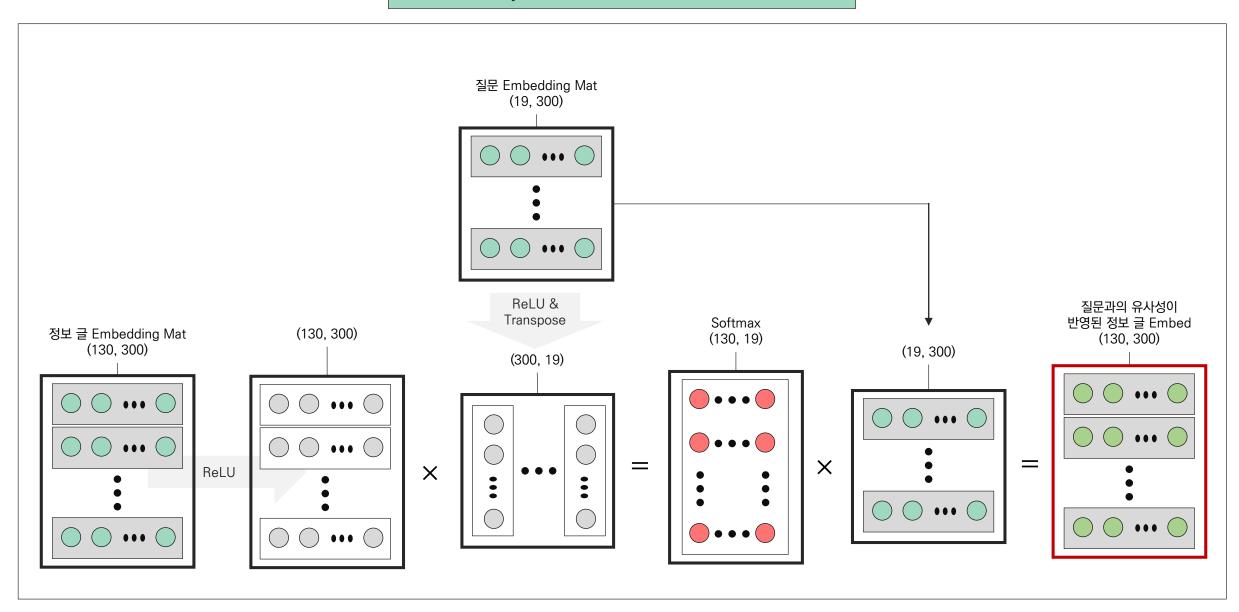
ToBigs 7th Conference - TMI 팀



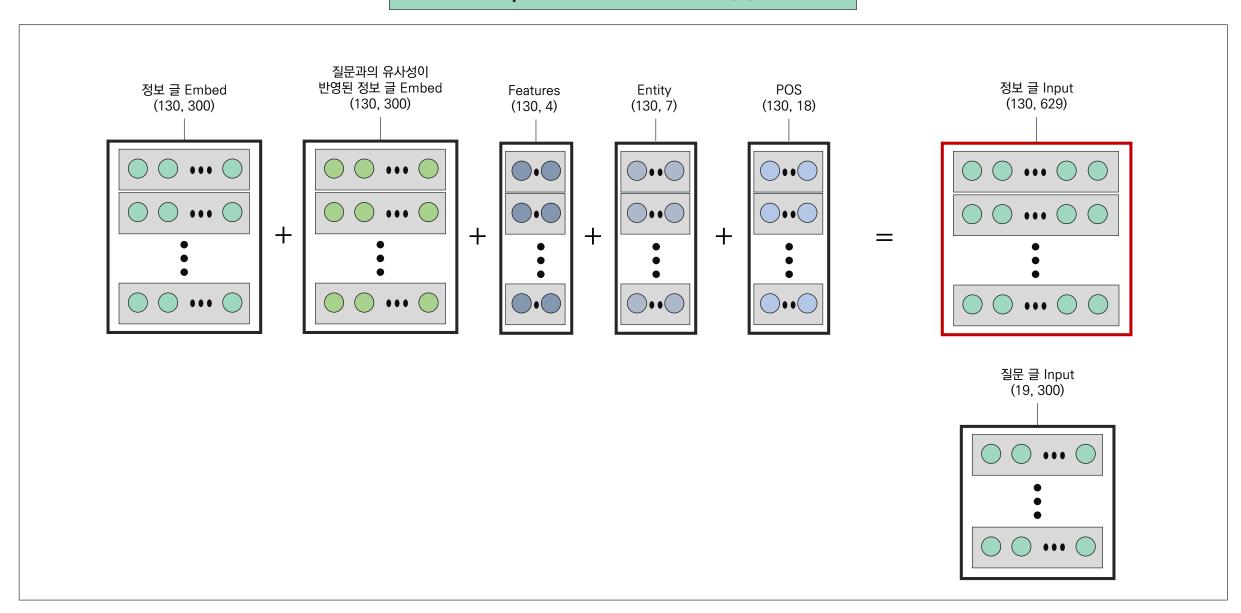
본 연구에서 적용한 DrQA 모델의 전체적인 구조는 아래와 같음



Input Data 준비 과정 (1)

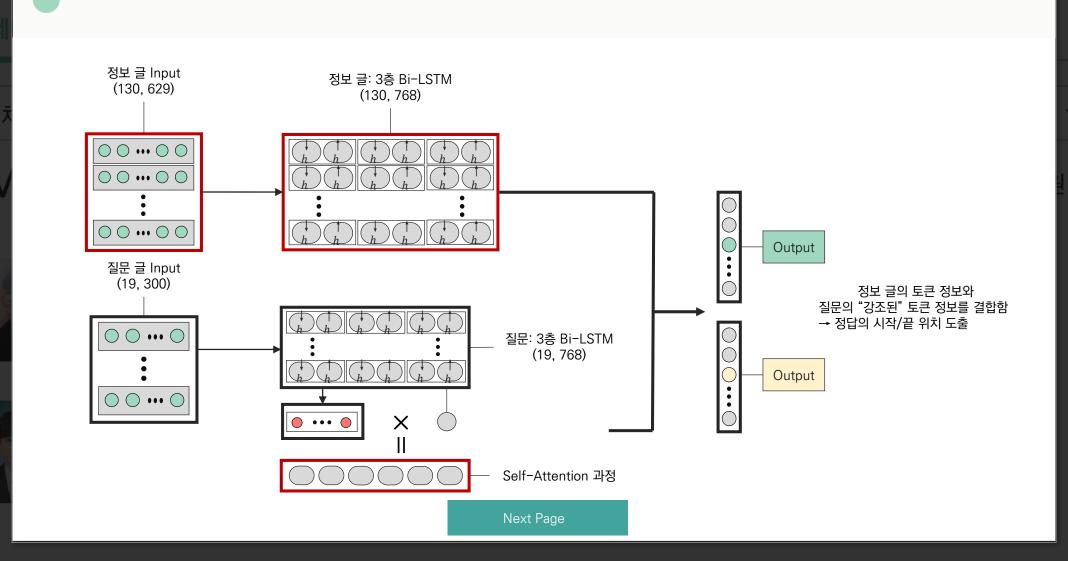


Input Data 준비 과정 (2)



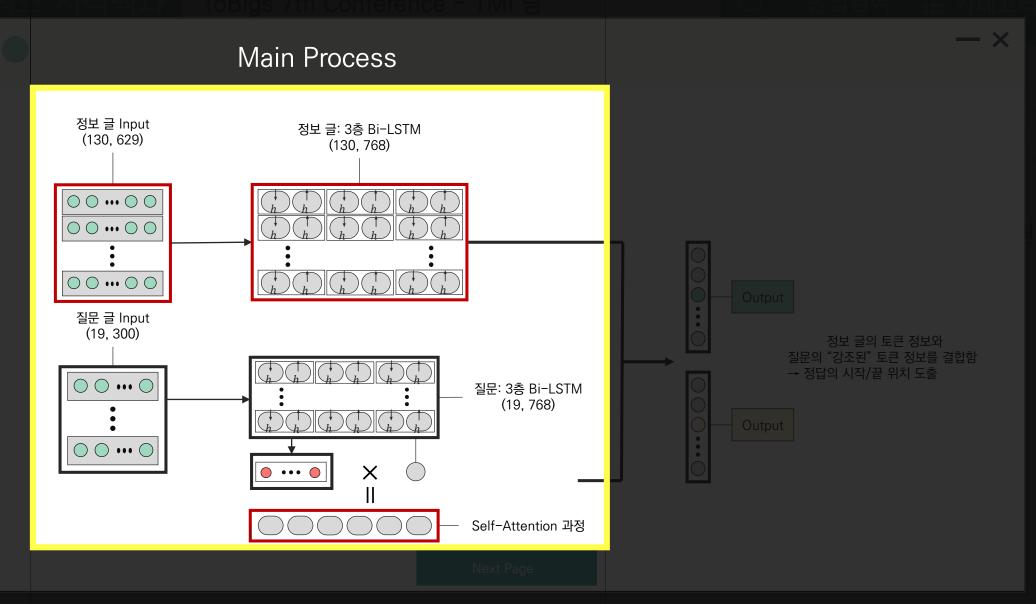
QA 모델 전체 구조

ToBigs 7th Conference - TMI 팀

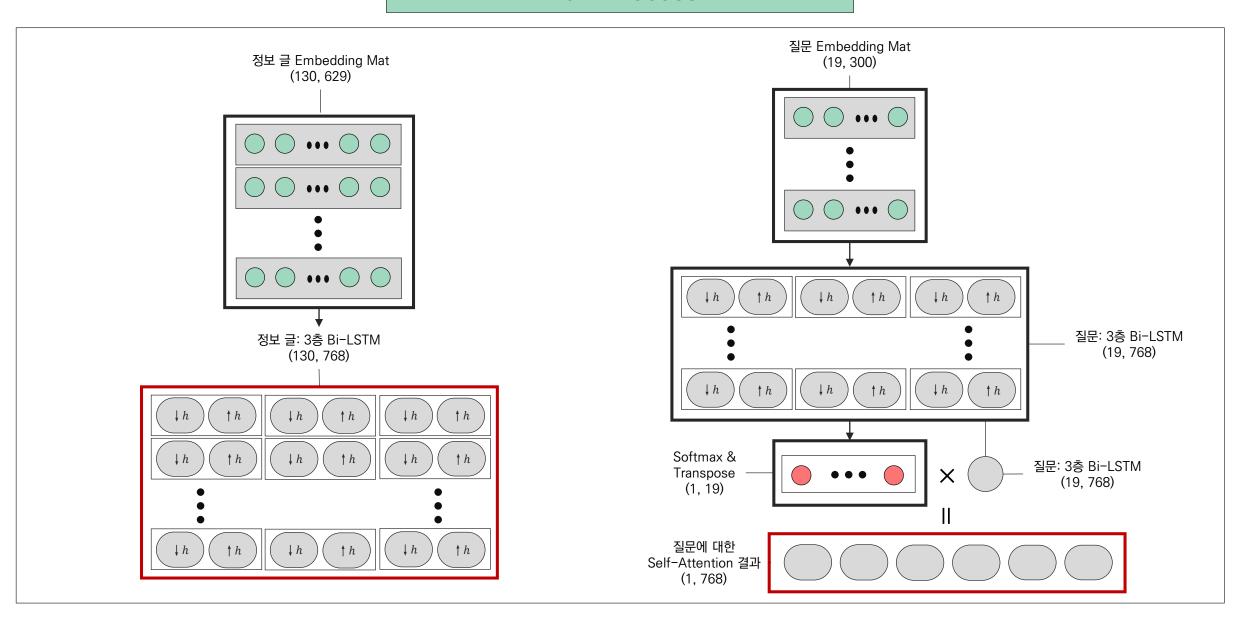


 $- \times$

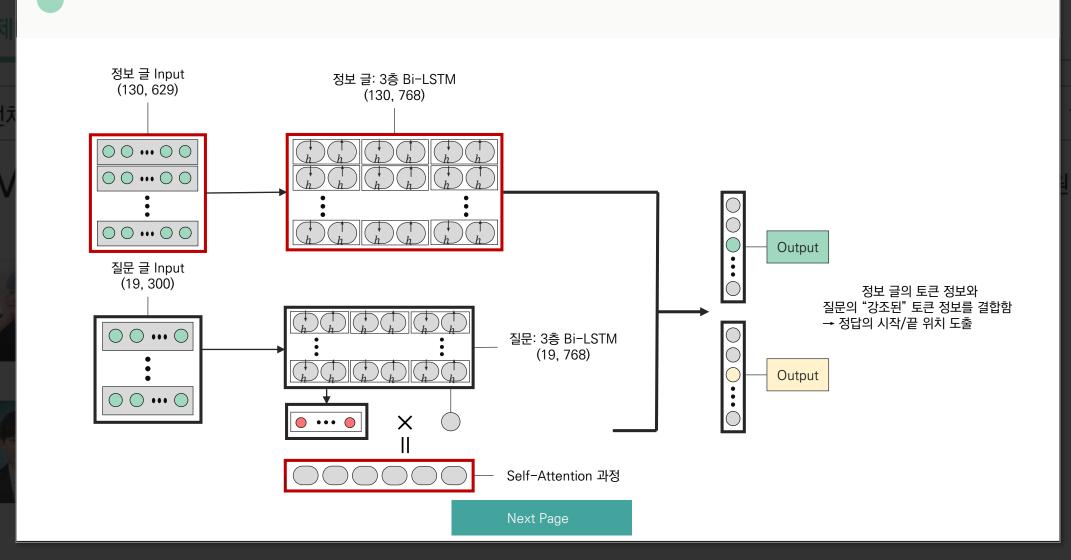
본 연구에서 적용한 DrQA 모델의 전체적인 구조는 아래와 같음



Main Process

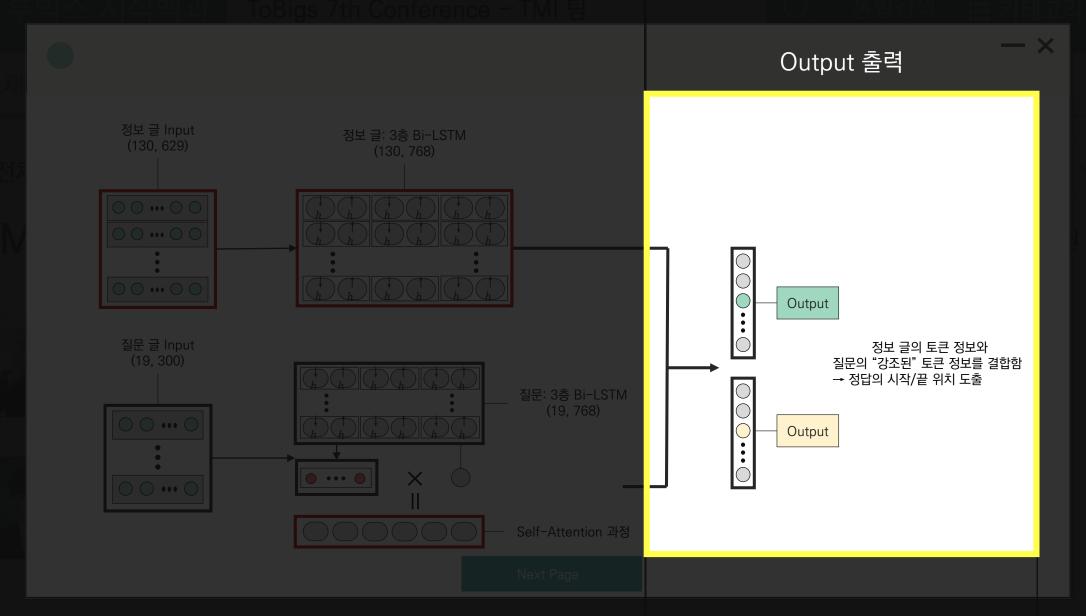


ToBigs 7th Conference - TMI 팀

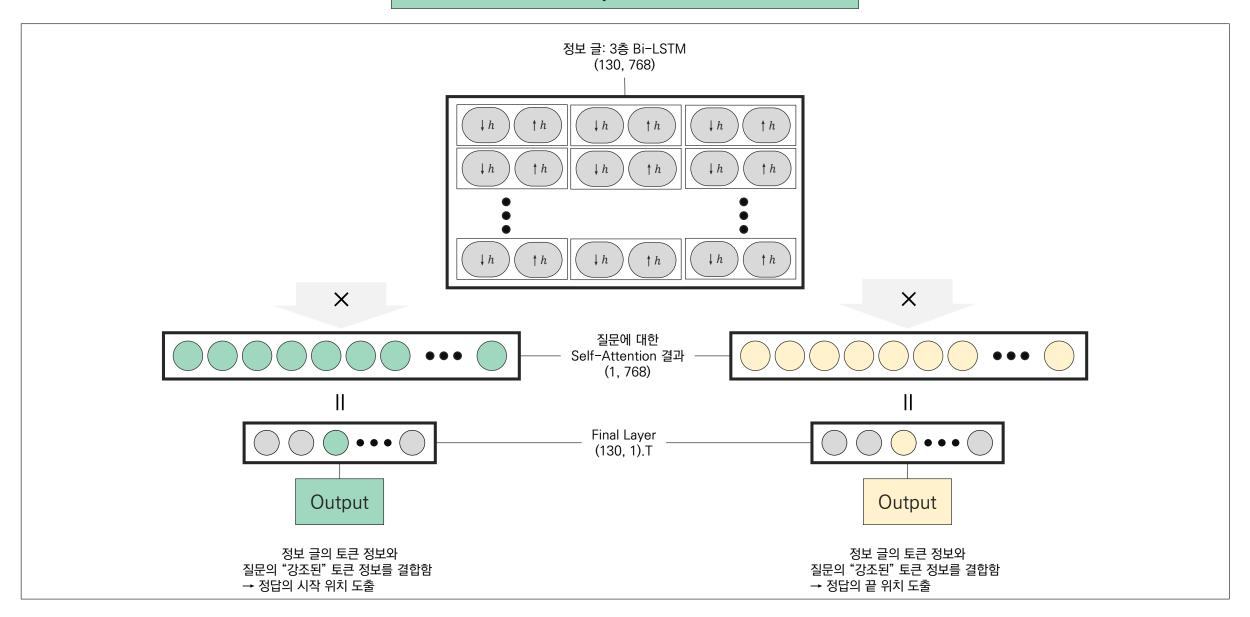


 $- \times$

본 연구에서 적용한 DrQA 모델의 전체적인 구조는 아래와 같음



Output 출력



Chapter		
Ch1. 주제 선정 배 Ch2. 모델 Overvi		
Ch3. 데이터 수집 Ch4. 모델 세부 구		
Ch5. 결과 및 QA5		
Ch6. 결론		
Date		
2 nd Round Place		

성능 평가

ToBigs 7th Conference - TMI 팀



EM, F1 Score

	EM	F1	
Basic	62.1	74.8	
NER 적용 후	64.2	77.1	

시연 예시

ToBigs 7th Conference - TMI 팀

 $- \times$

영국의 에브벨리, 프랑스의 엘라니크의 것과 규모의 스톤서클(環狀列石)의 유구(遺構)가 있는 것으로 유명하다. 지름 114m의 도랑과 도랑 안 전체쪽에 만들어진 제방에 둘러싸여 2중의 고리 모양으로 세워진 82개의 입석(立石)의 뽑힌 자리가 보인다. 중심부에는 2중으로 환상열석과 말발굽 모양의 열석이 둘러쳐 있다. 바깥쪽의 환상열석 은 지름이 30m인데 30개의 열석이 늘어서 있고, 그 위에 순석(楯石)을 난간처럼 걸쳐 놓았으며 지름 23m의 안쪽 열석에는 순석은 없다.

Question: 스톤헨지는 어느 나라의 유적일까

Answer: 영국

Evidence: 오로라는 태양에서 날아온 대전입자가 지구 자기장과 상호작용하여 극지방 상층 대기 모든 방향으로 내뿜는)이라 하는데, 태양에서 시구에 노날아는 내무문의 태양풍은 시구의 사기상 하지만 일부는 지구의 자기장에 끌려 반알렌대(Van Allen belt)라 불리는 알렌대는 조개 모양으로 지구 주위에 구부려져`있고 극쪽에서는 지표에 근접해서 구부려져 있다 . 대기 속에서 공기 분자와 대전입자가 서로 충돌하면 기체 분자 내부의 전자가 여기(excitatio n)된다. 그리고 원래 상태보다 높은 에너지를 가진 전자가 다시 원래 상태로 돌아오면서 빛을 방출한다.

Question: 태양풍이 뭐야

Answer: 태양 으로부터 모든 방향 으로 내뿜는 이런 플라스마 의 흐름

34 투빅스 TMI 팀 2019-01-12



Evidence: 조선 세종시대 최고의 과학자로 기억되는 장영실(蔣英寶), 역대 과학자 가운데 장영실만큼 이름이 회자되는 인물도 없지 않을까 싶다. 장영실은 자동으로 시간을 알려주는 물시계, 자격루를 한국 최초로 만든 인물이다. [조선왕조실록]에도 여러 번 능상할 성도로 유명인이었시만, 성작 그의 삶은 베일에 가려신 인물이기도 하다.장영실이 어떻게 출생하여 성장했는지 정확히 밝혀지지 않은 까닭은 그의 출생 배경에서 비롯된 것이다. 세종의 아낌없는 사랑을 받았다고 전해지는 장영실은 동래현의 관노(官奴), 즉 노비였다. [세종실록]에는 장영실의 부친은 원(元)나라 사람으로 소주(蘇州)·항주(杭州) 출신이고, 모친은 기녀였다고 전한다. Question: 물시계 자격루를 발명한 사람은

Answer: 장영실

Evidence: 닌텐도는 현재 소니컴퓨터엔터테인먼트(Sony Computer Entertainment), 세가(Sega) 안 더불어 일본을 대표하는 비디오 게일기 및 비디오 게임 제작업체이다. 원래 닌텐도는 화투나 트럼프를 만들던 교토[京都]의 전통 기업 기었으나, 1949년 가업을 계승한 야마우지 히로시가 새로운 장난감 개발로 눈을 돌디던지 오늘날의 게임기 업체로 변신하는 기반을 다졌다. 닌텐도는 1985년 선보인 게임기 패미콤(Famicom)과 1990년에 선보인 슈퍼패미콤(Super Famicom)을 통해 10년 가까이 전세계 게임 시장을 지배했으며, 닌텐도에서 제작한 슈퍼마리오브라더스는 비디오 게임의 상징물로 부상될 만큼 높은 인기를 누렸다.

Question: 닌텐도는 원래 무엇을 만드는 회사였나

Answer: 화투 나 트럼프

시연 예시

ToBigs 7th Conference - TMI 팀

Question :프레디 머큐리의 고향은

기 로파르츠는 파리음악원에서 T.뒤부아, 마스네, 프랑크에게 작곡과 대위법·화성법 등을 배움. 기 로파르츠는 주로 작곡활동에 전 념하였으나 뒤에 당시 및 스트라스부르의 음악학교 교수로 있으면서 P.O.페루 등, 후진양성에 힘씀. 기 로파르츠는 작품에는 고양 트 전설이나 민요를 지적이면서도 화려한, 프랑스적 음색으로 기악화한 것이 많음. 기 《 브르타뉴의 정경》, 오페라《 고향 Le Poys》, 그리고 약간의 가곡 등이 있음. 는 1926년 베이징으로 가서 베이징 사관학교에 입학함. 이육사는 루<u>쉰의 소설</u> 943년 중국으로 갔다가 귀국, 이 해 6월에 동대문경찰서 형사에게 체포되어 베이징으로 압송, 이듬해 는 1968년 고향인 경상북도 안동에 육사시비가 세워짐. 이육사는 베이징 유학 시절에 주더만는 독일의 자연주의 극작가. 지금 내림. 어린 주더만는 베를린 ·쾨니히스베르크 두 대학에서 공부함. 주더만는 《 우수부인 Frau Sorge》(1887)을 처녀작으로 발표하여 걸작임. 주더만는 그는 극작가로서도 한때 G.하우프트만과 견줄 유린당하고 버림받은 누이동생에게 돈으로 해결하려는 정조를 주더만는 4막극인 《 고향 Heimat》(1893)은 개인의 권리와 자유를 주장하는 묘사함. 주더만는 그의 희곡은 시대 사조를 교묘하게 혹평을 <mark>의어, 일</mark>부 비평가들의 받기도 함. 프레디 태어남. 프레디 머큐리는 1960년대 후반 아이벡스(Ibex), 사우워 활동암. 프레니 머큐리는 귄의 멤버로 《 어 나이트 앳 디 오페라》(1975년 영국 앨범차트 1위) 드(News Of The World)》(1977년 영국 앨범자트 4위), 《 더 게임》(1980년 미국 빌보드 앨범자트 1위) 등의 대표앨범을 《 미스터 배드 가이》(1985년 영국 앨범차트 6위), 1988년 스페인 출신의 소프라노 《 바르셀로나(Barcelona)》(1988년 영국 앨범차트 15위)를 연이어 발표함. 후천성면역결핍증(에이즈) 환자임을 스스로 밝힘. 프레디 머큐리는 날인 11월 24일 에이즈 합병증으로 `프레디´머큐리는 퀸의 멤버로 프로그레시브 록・글램 기여함. 프레디 머큐리는 4옥타브를 넘나드는 힘 있는 보컬과 발표한 《 어 나이트 앳 디 오페라》 《더게임》, 솔로로 발표한 비테즈슬라프 네즈발는 체코의 시인으로 1930년대 쉬르리얼리스트의 기수로 공부한 후 시작에 교사의 아들로 태어남. 비테즈슬라프 네즈발는 프라하대학에서 영향을 받아 1930년대에는 체코 쉬르리얼리스트의 기수가 되어 많은 작품을 직전에 그룹을 해산하였고, 파시즘에 격렬히 반대함. 비테즈슬라프 네즈발는 나치스 독일로부터 체코가 해방된 뒤에는 사회 비테즈슬라프 네즈발는 대표적인 시집으로는 《 에디슨 Edison》(1928) 《 평화의 노래》(1950) 《 고향》(1951) 등 이 있음. 비테즈슬라프 네즈발는 시집 외에 《 전위예술 이론》, 번역 시극 《 마농 레스코 Manon Lescaut》(1940), 미완성의 회상기 《 나의 생애에서》(1959) 등이 있음.

Answer: 탄자니아 잔지바르

ToBigs 7th Conference - TMI 팀

Question: 아인슈타인이 발표한 논문은 ["]르네 아르코스는 아베이파의 시인으로 유명함. 르네 아르코스는 《 외로프》지를 편집함. 르네 아르코스는 R.롤랑 등의 미술서 간행을 위해 사브리에 를 창설함. 르네 아르코스는 주요저서로 시집 《본질적 혼》(1898~1902), 《태어나는 것》(1910), 극시 《잃어버린 섬》(1913), 소설 《병영》(1921), 《타인》(1926) 등이 있음. 알베르트 아인슈타인는 독일 태생의 이론물리학자. 알베르트 아인슈타인는 1879년 3월 14일 독일 울름에서 유대인으로 알베르트 아인슈타인는 그의 아버지는 평범한 사업가였으며 가족들은 유대교 의식에 얽매이지 않음. 알베르트 아인슈타인는 어린시절 행동이 잦았다고 전해짐. 알베르트 아인슈타인는 1901년 대학을 졸업함. 있다는 브라운운동의 이론, 물리적 시공간에 대한 기존 입장을 완전히 뒤엎은 특수상대성이론 단 8수만에 작성된 것이지만 그동안의 인식을 신완시긴 논문으로 평가되었음. 일베르드 아인슈타인는 철악사상에도 - 몇 가지 뜻밖의 이론, 특히 질량과 에너지의 등가성의 발견은 원자폭탄의 가능성을 예언한 것이었음. 알베르트 아 빌헬름연구소 소장으로 취임함. 알베르트 아인슈타인는 1920년대에 됨. 알베르트 아인슈타인는 유대인 출신인 그는 유대민족주의 시오니즘운동의 지지자, 평화주의자로서 헬렌 프랑켄탈러는 미국의 제 2세대 추상표현주의 여성화가. 추상표현주의 화가인 한스 호프만(Hans Hofmann)에게서 그림을 배움. 헬렌 프랑켄탈러는 초벌칠이 되지 부어버림으로써 얼룩이 지거나 스며들도록 하는 이 기법은 화폭과 형상의 구분을 극복하여 완벽한 평면성에 이르도록 는 1953년 프랑켄탈러는 추상표현주의 이론을 정립한 예술 비평가 클레멘트 그린버그(Clement Greenberg)의 눈에 띄면서 뉴욕 미술계에 데뷔함. 헬렌 프랑켄탈러는 그는 1964년 그린버그가 기획한 '후기회화적 추상 Post-Painterly Abstraction' 전시에 참여하여 추상표현주의의 2세대 굳건히 함. 헬렌 프랑켄탈러는 그는 추상표현주의 동료화가인 로버트 머더웰(Robert Motherwell)과 결혼함. 헬렌 프랑켄탈러는 주요 작품에는 《 산괴 바다 Mountains and Sea》(1952), 《 호프부르크 궁전 Hofburg Palace》(1956), 《 왕복 여행 Round Trip》(1957), 《 문 Door》(1976~79) 등이 있음. 한스 호프만는 독일 태생의 미국 추상표현주의 화가이며 교육가. 한스 호프만는 1880년 독일 남부 바이에른주에 있는 바이센부르크에서 태어남. 한스 호프만는 1904년 파리에서 공부할 당시 마티스와 들로네, 피카소, 브라크 등을 알고 지냄. 한스 호프만는 1915년 뮌헨에 최초로 미술학교를 한스 호프만는 1932년 미국으로 이주한 후에도 캘리포니아대학교와 뉴욕의 아트 스튜던츠 리그에서 학생들을 가르침. 한스 호프만는 잭슨 폴록은 그에게서 절대적인 영향을 받음. 한스 호프만는 호프만의 야수파의 영향을 받은 재현적인 구상에서 출발함. 한스 호프만는 그는 예술을 위한 예술은 오직 순수한 추상이어야 한다는 철학 아래 순수 그려나감. 한스 호프만는 그는 교육에 너무 집중한 나머지 나이 60대 중반이 된 1944년이 되어서야 미국에서 처음으로 개인전을 는 직사각형의 색면들이 불규칙하게 쌓여 완성된 그의 작품은 유럽의 구성주의와 미국의 추상표현주의의 만남이라고도 할 는 1966년 미국 뉴욕에서 사망함. 한스 호프만는 주요 작품에는 《봄 Spring》(1940),《먹이 The Prey》(1956),《대양 The Ocean》(1957),《문 The 세쿤두스 Nulli Secundus》(1964) 등이 있음. 앨프리드'아인슈타인는 독일 출신의'미국의 음악사가로 모차르트와 이탈리 권위자임. 앨프리드 아인슈타인는 유명한 물리학자 A.아인슈타인의 육촌 동생임. 앨프리드 아인슈타인는 뮌헨대학교에서 음 Gate》(1959~60), 《 눌리 세쿤두스 Nulli Secundus》(1964) 등이 있음. 《 음악학잡지 Zeitschrift für Musikwissenschaft》를 편찬함. 앨프리드 아인슈타인는 모차르트와 이탈리아 마드리갈연 세계적 권위자인 그는 나치스의 손을 피해 1939년에 미국으로 건너가 미국에서 교편을 잡고, 1945년 미국에 귀화함. 앨프리드 아인슈타인는 L.쾨 모차르트 작품표의 제3판(1937)은 그의 저술이며, 《 모차르트-그의 생애와 작품》 《 슈베르트-음악적 초상》 Answer: 특수상대성이론 논문

ToBigs 7th Conference - TMI 팀

Question :최초로 달 착륙에 성공한 인물은

토마스 데만트는 1964년 독일 뮌헨에서 태어남. 토마스 데만트는 1989년부터 1992년까지 독일 뒤셀도르프 예술 아카데미를 다님. 토마스 데만트는 19 공부함. 토마스 데만트는 데만트의 사진은 정치적, 사회적, 역사적으로 모형으로 제작하여 이를 촬영하는 독특한 제작 방식을 가지고 있음. 토마스 모형을 만듦. 토마스 데만트는 "나는 소요해 붙임. 토마스 데만트는 오려서 만든 흔적만을 어렴풋이 데만트는 아이러니하게도 이것은 우리가 매일 접하는 신문과 인터넷을 뉴스, 있음을 보여줌. 토마스 데만트는 사진이 가진 전통적인 수 살롱 Salon》(1997), 《 터널 Tunnel》(1999), 《 유령 Ghost》(2003), 《 착륙 Landing》(2006), 《 프란츠 카프카는 유대계의 독일인 작가. 프란츠 카프카는 고등학교를 후에 법원에서 1년간 시보로 일함. 프란츠 카프카는 프라하에 있는 직장을 보험공사로 편집자가 된 M. 브로트를 알게 되어 문학에 대한 정열이 더욱 뜨거워져 프란츠 카프카는 그 사이 훗날 카프카 전집의 결혼 준비Hochzeitsvorbereitungen auf dem Lande》(1906) 등을 》(1925년 간행), 연말에 《 변신 Die Verwandlung》(1916년 간행)을 써서, 이 해는 최초의 중요한 결실기가 유대계 독일이라는 특이한 환경으로 인해 지냈는데, 받아, 여러 곳으로 요양을 하며 (1924년 간행)를 비롯한 단편을 성 Das Schloss》(1926년 간행) 《 배고 전전함. 프란츠 카프카는 그 장편소설 《 프란츠 카프카는 1924년 4월 예술가 Hunger Künstler》(1924년 간행)를 빈 교외의 우주비행사. 닐 암스트롱는 퍼듀대학교에서 항공학을 전공한 뒤, 출격하였으며 , 그후 미국항공우주국(NASA)에 암스트롱는 조종사로 78회 으즈비해의 하여 아제나 위성과 도킹에 성공함. 최초의

인간에 암스트롱는 '이것은 한 있어서는 작은 1969년 한국을 방문한 바 있음. 인물임. 가르느랭는 1769년 프랑스 파리에서 태어남. 가르느랭는 최초로 강하한 가르느랭는 세계에서 낙하산을 타고 군사적인 전략적 가르느랭는 그것의 헝가리 3년간 포로생활을 함. 가르느랭는 풀려난 후 1797년에 본인의 첫 부다에서 만든 낙하산을 직접 타고 내려오기 위해, 그해 가르느랭는 자신이 부력으로 상승하다가, 3200피트(약 975m)의 높이에서 낙하산과 기구를 는 기구의 방지하는 통풍구가 없었기 때문에, 하강하는 내내 크게 흔들렸지만, 다행히도 는 1799년에는 그의 아내인 잔-주느비에브 가르느랭(Jeanne-Genevieve Garnerin)이 여성으로서는 낙하산 강하 시범을 보임. 가르느랭는 1823년 8월 18일 파리에서 기구를 이규상는 1727년(영조 3)에 태어나 학행으로 천거되어 조산대부를 제수받은 이규상는 1799년(정조 23)에 죽었다는 것만 《 한산이씨세보》에 전해짐. 이규상는 말년에 18세기 조선의 내로라 하는 교류함. 알려지지 않은 채 1935년에 간행된 한산이씨의 문집인 《한산세고》에 묶여 있다가 최근 학계에 물지-병세재언록》이라는 제목으로 창작과비평사에서 단행본으로 출간되었음. 이규상는 조선 후기의 인물 전기로는 중요한 비중을 차지할 갖추고 있는 저작물로, 총 12권 7책으로 구성되어 있음. 이규상는 18세기 당대의 인물 180여 명이 수록되어 있어

지식체계를 엿볼 수 있음.

ToBigs 7th Conference - TMI 팀



| 태어남. 윤동주는 일본으로 건너가 도쿄에 있는 릿쿄대학 영문과에 입학하였다가(1942), 다시 도시샤대학 영문과 윤일주와 당숙인 윤영춘도 시인이었음. 윤동주는 그의 초간 시집은 하숙집 친구로 함께 지냈던 정병욱이 자필본을 첫 작품으로 <삶과 <쉽게 쓰여진 시> 등이 대표적인 그의 후기 1938년 모스크바로 돌아와 시를 쓰기 시작함. 카자케비치는 제2차 세계대전에는 의용군으로 모스크바 공방전, 베를린 공략전 참전하였으며, 전후에 소설가로 전향함. 카자케비치는 《 별 Zvezda》(1947) 《 오데르의 봄 Vesna na Odere》(1949) 등의 전쟁물을 쓰다가 후 암으로 죽었음. 카자케비치는 유작에`《 푸른 노트 Sinyaya tetrad'》(1961)가 경지를 열었으나, 얼마 러시아의 소설가 극작가. 보리스 라브레뇨프는 모스크바대학교 법과를 졸업하였으나 법관의 보리스 라브레뇨프는 국내전에서 취재한 중편소설 《 바람 Veter》(1924) 《 41번째의 사나이 Sorok pervyi》(192 얻은 에피소드를 엮은 것인데, 지나치게 도식화 유형화한 흠이 있음. 보리스 라브레뇨프는 다시 로 극작가로서의 지위를 확보함. 보리스 라브레뇨프는 《 흑해 수병의 노래》(1943) 《 아메리카의 소리 Golos Ameriki》(1949) 등의 작품을 그리펜베리는 국토 및 국민을 찬양하는 애국적 작품이 많음. 베르텔 그리펜베리는 와일드와 보들레르의 영향을 받은 흔적이 보임. 베르텔 그리펜베리는 주요 저서에는 《 시 Dikter》(1903), 《 위대한 길 Vid a vägar》(1904), 《 검은 소네트 Svarta sonetter》(1908), 《 타바스틀란드의 밤 Aftnar i Tavastland》(1911), 《 그림자 유희 Skuggspel》(1912) 등 클로드 시몽는 프랑스의 소설가로 누보 로망(반소설)의 대표적 작가임. 클로드 시몽는 마다가스카르섬 안타나나리보에서 태어남. 클로드 로망(반소설)의 대표적 작가임. 클로드 시몽는 기억이 의식으로 되살아나는 무질서하고 유동적인 형태를 그대로 본뜬, 있는 내적 독백의 문체를 구사하여, 자유에 대한 노스탤지어와 죽음으로의 접근에서 빚어지는 상극이 불러일으키는 내적 갈등을 멋있는 필치로 재현함 . 클로드 시몽는 1985년도 노벨문학상을 수상함. 클로드 시몽는 주요 작품에 A.카뮈의 《 이방인》과 비교되는 《 사기꾼 Le Tricheur》(1945), 《 바람 Le Vent》(1957), 《 초원 L'Herbe》(1958), 《 플랑드르로 가는 길 La Route des Flandres》(1960, 렉스프레스상 수상), 《 궁전 Le Palace》(1962) 등

Answer : 윤일주

ToBias 7th Conference - TMI 팀

문화지도위원 등에 종사하다가 1938년 모스크바로 돌아와 시를 쓰기 시작함. 카자케비치는 제2차 세계대전에는 의용군으 참전하였으며, 전후에 소설가로 전향함. 카자케비치는 《 별 Zvezda》(1947) 《 오데르의 한 얼마 후 암으로 죽었음. 《 광장의 봄》 (1956)으로 새로운 경지를 열었으나, 워런 버핏는 투자의 귀재라고 불리며 20세기를 대표하는 미국의 둘째 아들로 태어남. 워런 버핏는 11살 때에는 투자가의 때까지 펜실베이니아대학 와튼 비즈니스 스쿨, 네브래스카-링컨대학, 컬럼비아대학 Buffett Partnership Ltd)이라는 억만장자이면서도 검소한 생활태도를 지니고 있으며 2006년 재산의 생리·의학상을 수상한 오스트레일리아의 2005년 노벨 의학과를 거쳐, 1961년 같은 대학교 대학원을 르병원 임상병리학실에서 일한 뒤, 1968년부터 태어남. 워런는 애들레이드대학교 1967년까지 로열 멜버른병원 검사를 통해 위 아래쪽에 기생하는 굽은 형태의 작은 워런은 1979년 현미경 생체조직 사실을 있다 발견함. 워런는 당시만 해도 위에는 강한 있어 ∥ 보고했을 때 워런은 거짓말쟁이라는 비난을 받기도 함. 워런는 그는 더 나아가 일부 위암의 발생 원인은 물론, 항생제와 위산 분비 억제제를 사용하면 소화기관 정설이었기 때문에, 1982년 이 사실을 학회에 워런은 헬리코박터파일로리의 존재를 처음으로 제안한 공로로, 이를 입증한 마셜과 는 노벨상 외에 워런 앨퍼트상(1994), 호주의약협회 의학상(1995), 폴 에를리히상(1997) 등을 캘리포니아대학교 ·예일대학교 ·옥스퍼드대학교 등의 조교수로 근무하면서 계간지 《 서던 레ీ 《 서던 레뷰 The Southern Review》 확립함. 워런는 미네소타대학교 교수를 거쳐 1950년 이후 예일대학교 영문학 교수가 되었음. 워런는 현대 부르는 비평가들 중의 한 사람으로서 활약하여 치밀한 작품 한 작품이 많음. 워런는 '신비평파'라고 Understanding Poetry》(1938) 등 대학의 문학교육에 큰 영향을 끼친 텍스트의 편집을 담당함. 워런는 그밖에도 소설에 담배 Night Rider》(1939) 《 충분한 세계와 시간 World Enough and Time》(1950) 《 천사의 무리 Band of Angels》(1955) 《 동굴 Th 로맨스 Flood:a Romance of our Time》(1964) 등이 있음. 워런는 《 인종차별》 검사장 ·지사를 평가됨. 워런는 미국 제14대 대법원 수석판사. 워런는 캘리포니아주의 대법원 판사로 임명되었음. 워런는 이듬해에 공 사립학교에서 흑백아동을 차별하는 것은 마련함. 워런는 공립학교에서의 기도를 위헌이라고 판결하였으므로, 보수파로부터 "흑인을 학교로 끌어들이고 신을 내쫓았다" 선거구획 재편성에 관한 소송과 범죄혐의자의 인권사건에 적극적으로 관여하다가 1969년 노령으로 사임함. 워런는 194 공화당의 지명을 받고 부통령 후보로 출마하였으나 패함.

Answer: 오마하 의 현인

ToBigs 7th Conference - TMI 팀

Question :워런 버핏은 어떤 별명이 있나

Check Yourself!

Answer : 오바하 의 현인

Hey dude, Text Me Information!

ToBigs 7th Conference - TMI 팀



 $- \times$

- 1. 손을 들고 질문해 주세요.
- 2. 질문에 인물 이름, 작품 이름 (책, 음악, 시 등) 또는 날짜나 연도와 같은 구체적인 정보를 포함해주세요.
- 3. 작품 이름 등을 정확하게 입력해주시면 않으면 정답을 찾기 어렵습니다.
- 4. 특히 인물 이름은 정확히 말해 주셔야 합니다!
- 5. 지식백과 상에 없는 내용은 답할 수 없습니다.

질문 예시

프레디 머큐리의 대표 앨범은?

구약성서 중 <mark>잠언</mark>을 쓴 사람은?

Next Page

Hey dude, Text Me Information!

ToBigs 7th Conference - TMI 팀



Hint

영조: 속대전, 속오례의

마이클 잭슨: 오프 더 월로

이완용 : 황후폐하 치사문

안중근 : 동양평화론

모세: 출애굽기, 민수기

김구: 백천교중장비, 김주신도비

아돌프 히틀러 : 나의 투쟁

제갈량: 전출사표, 후출사표

원효: 대승기신론소, 금강삼매경론

마오쩌둥: 신민주주의론, 신단계론

메릴린 먼로: 아스팔트 정글, 나이아가라

아르키메데스: 평면의 균형에 대하여

이육사: 광야

조지 오웰: 동물농장

김소월: 진달래꽃, 엄마야 누나야

오드리 헵번: 로마의 휴일

윤봉길 : 농민독본

정몽주: 포은집

파블로 피카소 : 게르니카

솔로몬: 잠언, 아가

빈센트 반 고흐 : 별이 빛나는 밤

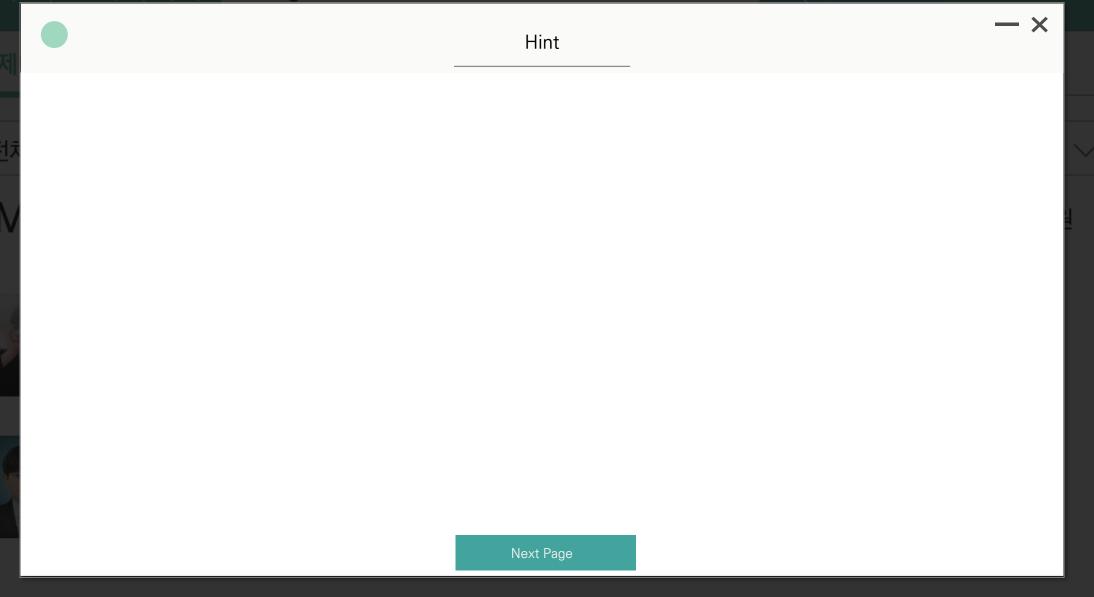
안데르센: 벌거숭이 임금님

Next Page

Ch5: 결과 및 QA봇 서비스 시연

Hey dude, Text Me Information!

ToBigs 7th Conference - TMI 팀



Hey dude, Text Me Information!

ToBigs 7th Conference - TMI 팀



Question :진달래꽃을 쓴 시인은

Answer : 김소월

Time: 0.3903s

Question :로마의 휴일 작품 주인공은

Answer : 오드리 헵번

Time: 0.1615s

Question :프레디 머큐리의 솔로곡은

Answer: 미스터 배드 가이

Time: 0.4602s

Question :유관순이 했던 운동은

Answer : 31 운동

Time: 0.4079s



Chapter		
Ch1. 주제 선정 배경		
Ch2. 모델 Overvie Ch3. 데이터 수집 및		
Ch4. 모델 세부 구3		
Ch5. 결과 및 QA봇	니연	
Ch6. 결론		
Date		
2 nd Round Place		

본 연구는 한국어판 QA 모델을 구축한 것에 있어 의의를 가지며, 세부적인 구조 향상을 꾀할 수 있다면 성능 개선을 기대할 수 있음



결과 및 결론

- Q&A 시스템은 여러 세부 TASK들이 합쳐진 종합 TASK
- 하나의 질문으로 문서가 아닌 답을 찾아주는 Q&A 시스템 구현
- 학습에 필요한 임베딩 직접 학습
- 질문의 중요 토큰을 문서의 토큰과 매칭
- 한국형 SQuAD 공개로 정제데이터 활용

향후 연구 방향

- 토크나이저 성능 개선이 필요
- 문서유사도, NER, POS, 임베딩 각각이 하나의 큰 TASK
- 정보 글과 다른 형태의 질문 이해
- 카카오톡 플러스 친구 응답 대기 제한 시간 내 속도 개선
- 좋은 성능을 보이는 BERT 시도

Next Page



투빅스 지식백과

ToBigs 7th Conference - TMI 팀



통합검색

፧਼ 카테고리 보기 ▼

표제어 20

이미지 19

동영상 01

리스트 12

전체 20

어린이백과 2

학생백과 6

전체보기 🗸

TMI 팀 Text Me Information의 준말로 나이 합은 157 키 합은 1000 지금 입고 있는 옷의 가격 합은 230만원



김동규 | 투빅스 9기 국민대학교 빅데이터경영통계학과 # 최근에 아이패드 펜을 삼



박규리 | 투빅스 10기 경희대학교 경영학과 # 경희대 최대 주정뱅이, 노숙취



신훈철 | 투빅스 10기 홍익대학교 산업공학과 # 파마했는데 아무도 모름



장유영 | 투빅스 10기 고려대학교 통계학과 # 세영이 친구와 똑같이 생김



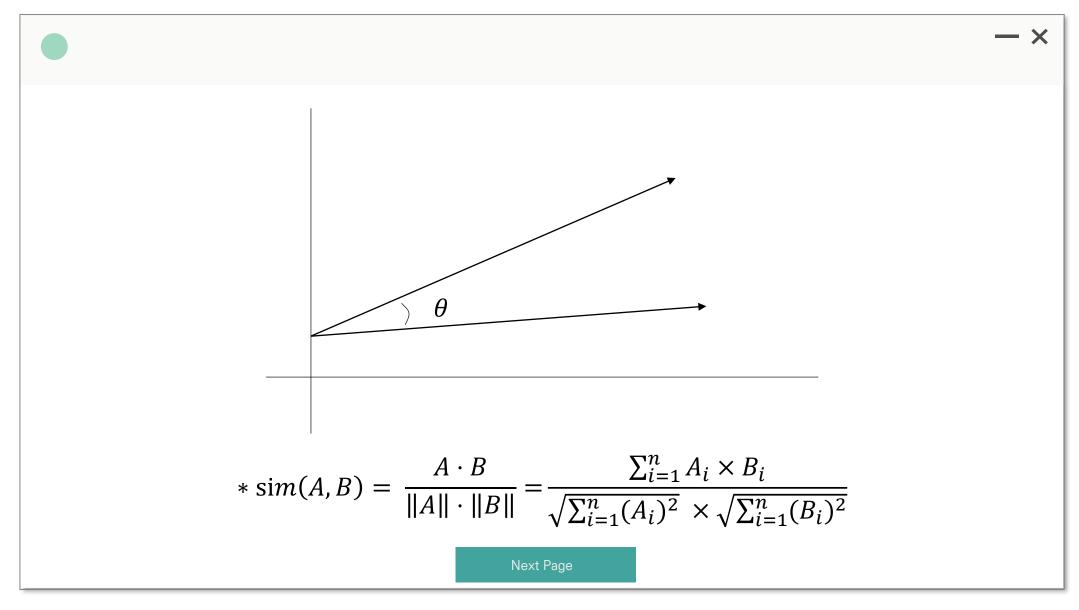
장청아 | 투빅스 10기 동국대학교 식품산업관리학과 # 선인장 다육이 엄마



최세영 | 투빅스 10기 덕성여자대학교 컴퓨터학과 # 고양이 유료 팬클럽 멤버



Appendix (1) - 코사인 유사도



Appendix (2) – Reference

 $- \times$

Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532–1543).

Chen, D., Bolton, J., & Manning, C. D. (2016). A thorough examination of the cnn/daily mail reading comprehension task. *arXiv* preprint arXiv:1606.02858.

Chen, D., Fisch, A., Weston, J., & Bordes, A. (2017). Reading wikipedia to answer open-domain questions. *arXiv preprint arXiv:1704.00051*.

https://github.com/hitvoice/DrQA & https://github.com/eagle705/Korean_NER_CNN_BiLSTM

https://medium.com/@florijan.stamenkovic_99541/rnn-language-modelling-with-pytorch-packed-batching-and-tied-weights-9d8952db35a9

https://terms.naver.com/list.nhn?cid=40942&categoryId=40942

https://korquad.github.io/ & http://www.aihub.or.kr/

https://terms.naver.com/list.nhn?cid=40942&categoryId=409425

Next Page