수능 국어 푸는 Al

김현수

ABOUT ME

- 통계학과 3 학년 학부생 → "부캠 수료"가 가장 큰 목표 (취업 x)
- 가장 자신있는 것은 "아이디어" \rightarrow 수능 국어 푸는 AI ? / 의견 내는 것을 좋아함
- 두 번의 학부연구생 → 인공지능공학과 (시계열 예측), 수학과 (Multimodal)
- **다양한 도메인에 관심 많음** → 금융, 음악, 패션, 법률, 의학 등 사실상 모든 분야 환영
- 경기도 광명시 거주 \rightarrow \geq / 오프 둘 다 가능 !

TMI

수능 국어 푸는 AI

김현수

수능 국어 푸는 AI?

Transformer NLP Natural Language Processing

BER Natural Language
Understanding

ural Language
derstanding

Natural Languag
Generation

GPT

(Natural Language Understanding)

(Natural Language Generation)

Transformer

NLP

Natural Language Processing

그래서 얼마나 잘 하는데?

NLU

Natural Language
Understanding

NLU

Generation

GPT

(Natural Language Understanding)

(Natural Language Generation)

placed with a single mask token. BART is particularly effective when fine tuned for text generation but also works well for comprehension tasks. It matches the performance of RoBERTa with comparable training resources on GLUE and SQuAD, achieves new stateof-the-art results on a range of abstractive dialogue, question answering, and summarization tasks, with gains of up to 6 ROUGE. BART also provides a 1.1 BLEU increase over a back-translation system for machine translation, with only target language pretraining. We also report ablation experiments that replicate other pretraining schemes within the BART framework, to better measure which factors most influence end-task performance.

[출처] BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, 2020.

6 ROUGE?

1.1 BLEU?

(BiLingual Evaluation Understudy score)

BLEU는 기계 번역 결과와 사람이 직접 번역한 결과가 얼마나 유사한지 비교하여 번역에 대한 성능을 측정하는 방법 → n-gram에 기반한 방법으로 언어에 구애받지 않고 사용할 수 있으며, 계산 속도가 빠름 (BLEU는 PPL과는 달리 높을 수록 성능이 더 좋음을 의미)

(Recall-Oriented Understudy for Gisting Evaluation)

ROUGE는 텍스트 자동 요약, 기계 번역 등 자연어 생성 모델의 성능을 평가하기 위한 지표이며 → 모델이 생성한 요약본 혹은 번역본을 사람이 미리 만들어 놓은 참조본과 대조해 성능 점수를 계산

ROUGE에서 Precision과 Recall은 아주 단순한 방법을 통해 구할 수 있다

시스템 요약 (모델 생성 요약):

the cat was found under the bed

참조 요약 (Gold standard, 대개 사람이 직접 만든 요약):

the cat was under the bed

- Precision

$$\frac{\textit{Number_of_overlapped_words}}{\textit{Total_words_in_reference_summary}} \quad \textit{Recall} = \frac{6}{6} = 1.0$$

- Recall

$$\frac{Number_of_overlapped_words}{Total_words_in_system_summary} \qquad Precision = \frac{6}{7} = 0.86$$

ROUGE-N, ROUGE-L, ROUGE-W, ROUGE-S, ROUGE-SU, ROUGE-N-precision, ROUGE-N-f1 ...

RDASS

Reference and Document Aware Semantic Evaluation Methods for Korean Language Summarization

Dongyub Lee¹, Myeongcheol Shin², Taesun Whang³, Seungwoo Cho², Byeongil Ko², Daniel Lee², Eunggyun Kim², Jaechoon Jo^{4*}

¹Kakao Corp.

jude.lee@kakaocorp.com

²Kakao Enterprise Corp.

{index.sh, john.w, kobi.k, daniel.e, json.ng}@kakaoenterprise.com

3Korea University, South Korea

{taesunwhang}@korea.ac.kr

⁴Hanshin University, South Korea

{ jaechoon }@hs.ac.kr

Abstract

Text summarization refers to the process that generates a shorter form of text from the source document preserving salient information. Many existing works for text summarization are generally evaluated by using recall-oriented understudy for gisting evaluation (ROUGE) scores. However, as ROUGE scores are computed based on n-gram overlap, they do not reflect semantic meaning correspondences between generated and reference summaries. Because Korean is an agglutinative language that combines various morphemes into a word that express several meanings, ROUGE is not suitable for Korean summarization. In this paper, we propose evaluation metrics that reflect semantic meanings of a reference summary and the original document, *Reference and Document Aware Semantic Score (RDASS)*. We then propose a method for improving the correlation of the metrics with human judgment. Evaluation results show that the correlation with human judgment is significantly higher for our evaluation metrics than for ROUGE scores.

RDASS는 〈본문, 정답 요약 문장, 예측 요약 문장〉 세 개의 관계를 동시에 고려하도록 고안됐습니다. RDASS도 0~1 사이의 숫자 값을 가지며 숫자가 높을수록 더 높은 성능을 가리킵니다. 계산 과정([그림 3])은 다음과 같습니다.

사전훈련과 미세조정을 모두 마친 버전인 FWA-SBERT를 이용해 문서(d), 정답 요약 문장(r), 모델이 생성한 요약 문장(p) 각각의 벡터 V_d , V_p , V_r 를 추출합니다. 그런 뒤, $\langle V_p, V_r \rangle$ 와 $\langle V_p, V_d \rangle$ 각각의 코사인 유사도 cosine similarity 10 인 $_{S}(p,r)$, $_{S}(p,d)$ 를 계산합니다. 마지막으로, 두 유사도의 평균값을 냅니다.

$$s(p, r) = cos(v_p, v_r) = \frac{v_p^{\mathrm{T}} \cdot v_r}{\|v_p\| \|v_r\|}$$

$$s(p,d) = cos(v_p, v_d) = rac{v_p^{\mathrm{T}} \cdot v_d}{\left\|v_p\right\| \left\|v_d\right\|}$$

$$RDASS = \frac{s(p,r) + s(p,d)}{2}$$

문서

리오넬 메시(30 · fc바르셀로나)가 자신의 <mark>서른 번째 생일을 가족과 함께 오붓하게 보냈다</mark>. 지난 24일 만 서른 살이된 메시는 자신의 인스타그램에 집에서 가족들과 함께 보낸 생일상을 찍은 사진을 올렸다. 메시는 오랜 그의 여자친구이자, 이제 아내가 되는 안토넬라 로쿠조(29), 아들 티아고가 함께 다정하게 사진을 찍었다.

정답 (요약) 문장

메시가 30번째 생일 함께한 이는 아내와 아들

R-1 R-2 R-L

모델이 생성한 문장

메시 30번째 생일, 가족과 함께 오붓하게 보내

0.14 0.00 0.14

문서

'슬기로운 의사생활'이 또다시 최고 시청률을 경신하며 고공행진을 이어갔다. 26일 방송된 tvN 2020 목요 스페셜 '슬기로운 의사생활' 3회는 케이블, IPTV, 위성을 통합한 유료플랫폼에서 가구 평균 8.6%, 최고 10%의 시청률을 기록했다. 3주 연속 시청률 상승세다.

정답 (요약) 문장 '슬기로운 의사생활' 최고 시청률 10% 돌파… 3회 연속 상승세 R-1 R-2 R-L
모델이 잘못 생성한 문장 '슬기로운 의사생활' 최저 시청률 10% 돌파… 3회 연속 하락세 0.78 0.63 0.78
모델이 바르게 생성한 문장 '슬기로운 의사생활' 최고 시청률 경신… 3주 연속 상승세 0.71 0.53 0.71

RDASS

Reference and Document Aware Semantic Evaluation Methods for **Korean Language Summarization**

Dongyub Lee¹, Myeongcheol Shin², Taesun Whang³, Seungwoo Cho², Byeongil Ko², Daniel Lee², Eunggyun Kim², Jaechoon Jo⁴*

ROUGE is not suitable for Korean summarization. In this paper, we propose evaluation metrics that reflect semantic meanings of a reference summary and the original document, Reference and Document Aware Semantic Score (RDASS). We then propose a method for improving 가지며 숫자가 높을수록 더 높은 성능을 가리킵니다. 계산 과정([그림 3])은 다음과 같습니다.

 V_a, V_b, V_r 를 추출합니다. 그런 뒤, $\langle V_b, V_r \rangle$ 와 $\langle V_p, V_d \rangle$ 각각의 코사인 유사도 cosine similarity 10 인 s(p, r), s(p, d)를 계산합니다. 마

$$s(p, r) = cos(v_p, v_r) = \frac{v_p^{\mathrm{T}} \cdot v_r}{\|v_p\| \|v_p\|}$$

그래서 얼마나 잘 하는데?

$$RDASS = \frac{s(p, \tau) + s(p, a)}{2}$$

리오넬 메시(30·fc바르셀로나)가 자신의 서른 번째 생일을 가족과 함께 오붓하게 보냈다. 지난 24일 만 서른 살이 된 메시는 자신의 인스타그램에 집에서 가족들과 함께 보낸 생일상을 찍은 사진을 올렸다. 메시는 오랜 그의 여자친 구이자, 이제 아내가 되는 안토넬라 로쿠조(29), 아들 티아고가 함께 다정하게 사진을 찍었다.

정답 (요약) 문장 메시가 30번째 생일 함께한 이는 아내와 아들

모델이 생성한 문장 메시 30번째 생일, 가족과 함께 오붓하게 보내

0.14 0.00 0.14

문서

'슬기로운 의사생활'이 또다시 최고 시청률을 경신하며 고공행진을 이어갔다. 26일 방송된 tvN 2020 목요 스페셜 '슬기로운 의사생활' 3회는 케이블, IPTV, 위성을 통합한 유료플랫폼에서 가구 평균 8.6%, 최고 10%의 시청률을 기록했다. 3주 연속 시청률 상승세다.

정답 (요약) 문장 '슬기로	약) 문장 '슬기로운 의사생활' 최고 시청률 10% 돌파… 3회 연속 상승세		R-2	R-L
모델이 잘못 생성한 문장	'슬기로운 의사생활' 최저 시청률 10% 돌파… 3회 연속 하락세	0.78	0.63	0.78
모델이 바르게 생성한 문장	'슬기로운 의사생활' 최 <mark>고</mark> 시청률 <mark>경신… 3주 연속 상승세</mark>	0.71	0.53	0.71

뤼이드는 이번 연구를 통해 추후 학습자와 쌍방향으로 대화할 수 있는 AI 튜터 개발의 기술적 기 반을 마련했다.

NAACL을 통해 공개될 뤼이드의 또다른 논문(논문명: GRAM: Fast Fine-tuning of Pre-trained Language Models for Content-based Collaborative Filtering)은 콘텐츠 기반의 지식 추적 모델 (Knowledge Tracing)을 보다 효율적으로 훈련시킬 수 있는 새로운 알고리즘을 제안했다.

과거에도 유저의 정오답을 예측하는 KT 모델에 콘텐츠를 활용하려는 시도는 있었지만, 지나치게 연산량이 많고 방대한 시간이 소요되어 실제 솔루션에 적용하는데 한계가 존재했다. 하지만 뤼 이드가 이번에 제안한 새로운 프레임 워크를 적용할 경우 특정 도메인에서 최대 150분의 1까지 훈련 시간을 단축시킬 수 있는 것으로 나타났다.

또 뤼이드가 제시한 모델은 유저 풀이 기록이 전혀 없는 신규 문제들에 대해서도 강력한 성· 상을 보이며 기존 모델 대비 정오답 예측 정확도(AUC)를 약 40% 향상시키는 것으로 나타났다.

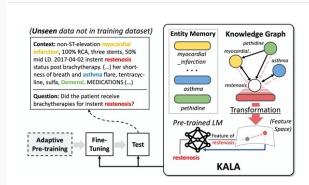
에이아이트릭스, 의료분야 자연어 처리 향상 언어모델 개발

뤼이드는 논문에서 제안한 콘텐츠 기반 KT 모델을 바탕으로 5월 중 산타 플랫폼에서 A/B 테스팅 을 통해 유저 데이터를 다각도로 검증한 후, 산타 제품에 바로 적용할 예정이다. 매월 산타 플랫폼 에 새로운 문제가 추가되고 있는 만큼, 유저

HOME > 뉴스 > ICT용복합

△ 최아름 기자 │ ② 승인 2022.08.23 14:47 │ ② 댓글 0

○ 정희석 기자 () 승인 2022,05,23 13:28 ○ 댓글 0



[라포르시안] 인공지능(AI) 기술 전문기업 에이아이트릭스(대표이사 김광준 유진규)가 국제 자 연어 처리(NLP) 학회 'NAACL 2022'(북미 전산 언어학 학회)에 참가해 새롭게 개발된 언어 모 델 논문을 발표한다고 23일 밝혔다.

NAACL은 ACL(Association for Computational Linguistics) EMNLP(Empirical Methods in Natural Lan guage Processing)와 함께 자연어 처리 분야에서 최고 수준의 국제학술대회로 꼽힌다.

에이아이트릭스는 미국 워싱턴주 시애틀에서 오는 7월 10일부터 15일까지 진행되는 NAAC L 2022에 참가해 '지식 증강 언어 모델 적응'이란 논문을 발표할 예정이다.

해당 논문에서 선보인 'KALA 프레임워크'는 의료분야에서 요구되는 전자의무기록에서의 질의 응답과 의료 논문에서의 질병 개체명 인식을 포함한 다양한 분야의 자연어 처리 태스크들에 서 기존 언어 모델 대비 더 높은 성능을 보인다.

기존 BERT와 같은 사전 학습된 언어 모델들은 기계 독해나 개체명 인식과 같은 다양한 태스크 🖁 보이고 있으나 의료와 같은 특수한 분야에서는 자연어 처리 성능이 더 욱 개선될 필요가 있었다. 이를 해결하기 위해 여러 방법론들이 제시됐지만 여전히 많은 계 산 비용을 필요로 한다는 단점이 있었다.



1. 제주 이어 강원도에 영리

2. 공공병원 쥐어짜고, 영리

병원 허용 추진...건강관

자-의료진 다 잃었다...존

의료민영화 정책"

리 민영화까지

폐 고민 내몰려

병원 허용 추진..."윤정부,

+ - ☑ 昼 력한 학습 경험을 제공할 수 있을 것으로 기대

일력 2022.08.29 09:20

한국특허정보원(원장 강경호, 이하 '정보원')은 지식재산(IP) 분야에 특화된 고성능 인공

정보원은 지난해 AI 언어 모델 KorPatBERT를 공개한 바 있다. 올해 새롭게 보급하는 KorPatELECTRA는 Google社에서 기존 언어모델 BERT 보다 학습 효율과 성능을 향상 시키기 위해 개발한 최신 언어모델 ELECTRA를 기반으로 구축한 언어모델이다 KorPatELECTRA는 특허분야 한국어 자연어처리에 적합하도록 국내 특허문헌 약 466만 건에 포함된 5억4천만 문장을 학습하여 구축하였다.

기존 언어모델과의 성능 비교 결과 KorPatELECTRA는 개체명 인식과 기계독해 분야에서 뛰어난 성능을 보였으며, 이를 통해 지능형 IP정보서비스 구축 시 언어모델 선택의 폭을 넓혀 줄 수 있을 것으로 기대된다.

신규 언어모델의 상세한 구축 과정과 학술적 성과는 추가적인 연구로 확장될 수 있도록 논

네이버클라우드, '클로바 스튜디오'에 튜닝 기능 추가 문장 요약, 문장 생성 등 다양한 NLP 과제 최적화 Q# 1 Q 1 UPDATE 101버클라우드. 로바 성능 극대화하는 itudio 튜닝 기능 추가

NAVER CLOUD PLATFORM

le) AI 플랫폼 '클로바 스튜디오(CLOVA Studio)'에 튜닝 기능

른 지식이 없이도 AI의 적용 가능성을 탐색하고 실제 서비스

·랫폼이다. 비개발자도 자유롭게 텍스트 기반으로 AI를 실험

른 사용자들이 작업한 AI를 둘러보고 활용할 수 있는 '익스플

:답을 나누는 '포럼' 공간으로 구성돼 있다.

"초대규모 AI 언어모델 성능 극대화해 누구나 AI 개발"

켈 최적화 기능이 강화됐다.

(*) (*) (*) 안전지키미로 시작하세요! 인기기사

한국정보통신공사협회 안전기술원

1 KT-SKT-LGU+-삼성SDS 등 대기업, 10억 미만

기업은행, 방송 제작시스템 계약 부당처리 ..

위험성 평가에서 대책마련까지 한번에

[KES2022] 공공 영상·음향·조명설비 사업 이.. 4 [기자수첩]미래 유산 '안전' 디지털 기술 옷 ...

5 "승계 세부담 줄여 투자 고용 활성화해야"

6 구내통신인프라 고도화 제도개선 속도낸다

7 건축물 인증서 관리, ICT 설비 인수과정 허술

8 우수 중기 제품 시장진출 지원 과제 성과 '부.

9 2050년 미래 한국, 경찰드론 로봇이 24시간 ..

10 [KES2022] 전자·정보통신산업 ESG 선점 위한.

최신기사

과기정통부, '2022 코리아 메타버스 어워드' 수 상작 9건 선정

다양성과 포용성을 위한 KAIST 선언문' 공표

지능 언어모델인 KorPatELECTRA를 확대 보급하겠다고 밝혔다.

문발표를 통해 내용을 공유하였고, 연구의 우수성을 인정받아 관련 저널에 게재 되었다.

홀수형

[1~3] 다음 글을 읽고 물음에 답하시오.

제 1 교시

어떤 독서 이론도 이 한 장의 사진만큼 독서의 위대함을

분명하게 말해 주지 못할 것이다. 사진은 제2차 세계 대전 당시 처참 하게 무너져 내린 런던의 한 건물 모습이다. ① 폐허 속에서도 사람 들이 책을 찾아 서가 앞에 선 이유는 무엇일까? 이들은 갑작스레 닥친



상황에서 독서를 통해 무언가를 구하고자 했을 것이다.

독서는 자신을 살피고 돌아볼 계기를 제공함으로써 어떻게 살 것인가의 문제를 생각하게 한다. 책은 인류의 지혜와 경험이 담겨 있는 문화유산이며, 독서는 인류와의 만남이자 끝없는 대화 이다. 독자의 경험과 책에 담긴 수많은 경험들의 만남은 성찰의 기회를 제공함으로써 독자의 내면을 성장시켜 삹을 바꾼다. 이런 의미에서 독서는 자기 성찰의 행위이며, 성찰의 시간은 깊이 사색하고 스스로에게 질문을 던지는 시간이어야 한다. 이들이 책을 찾은 것도 혼란스러운 현실을 외면하려 한 것이 아니라 자신의 삶에 대한 숙고의 시간이 필요했기 때문이다.

또한 ①독서는 자신을 둘러싼 현실을 올바로 인식하고 당면한 문제를 해결할 논리와 힘을 지니게 한다. 책은 세상에 대한 안목을 키우는 데 필요한 지식을 담고 있으며, 독서는 그 지식을 얻는 과정이다. 독자의 생각과 오랜 세월 축적된 지식의 만남은 독자 에게 올바른 식견을 갖추고 당면한 문제를 해결할 방법을 모색 하도록 함으로써 세상을 바꾼다. 세상을 변화시킬 동력을 얻는 이 시간은 책에 있는 정보를 이해하는 데 그치는 것이 아니라 그 정보가 자신의 관점에서 문제를 해결할 수 있는 타당한 정 보인지를 판단하고 분석하는 시간이어야 한다. 서가 앞에 선 사람들도 시대적 과제를 해결할 실마리를 책에서 찾으려 했던 것이다.

독서는 자기 내면으로의 여행이며 외부 세계로의 확장이다. 폐허 속에서도 책을 찾은 사람들은 독서가 지닌 힘을 알고, 자신과 현실에 대한 이해를 구하고자 책과의 대화를 시도하고 있었던 것이다.

- 1. 윗글을 바탕으로 할 때, ③의 답으로 적절하지 않은 것은?
- ① 인류의 지혜와 경험을 배우기 위해
- ② 현실로부터 도피할 방법을 구하기 위해
- ③ 시대적 과제를 해결할 실마리를 찾기 위해
- ④ 자신의 삶에 대해 숙고할 시간을 갖기 위해
- ⑤ 세상에 대한 안목을 키우는 지식을 얻기 위해

2. <보기>는 C과 같이 독서하기 위해 학생이 찾은 독서 방법이다. 이에 대한 반응으로 적절하지 않은 것은? [3점]

<보 기>—

해결하려는 문제와 관련하여 관점이 다른 책들을 함께 읽는 것은 해법을 찾는 한 방법이다. 먼저 문제가 무엇인지를 명확히 하고, 이와 관련된 서로 다른 관점의 책을 찾는다. 책을 읽을 때는 자신의 관점에서 각 관점들을 비교·대조 하면서 정보의 타당성을 비판적으로 검토하고 평가한 내용을 통합한다. 이를 통해 문제를 다각적 · 심층적으로 이해하게 됨으로써 자신의 관점을 분명히 하고, 나아가 생각을 발전시켜 관점을 재구성하게 됨으로써 해법을 찾을 수 있다.

- ① 읽을 책을 선택하기 전에 해결하려는 문제가 무엇인지를 명확 하게 인식해야겠군.
- ② 서로 다른 관점을 비교·대조하면서 검토함으로써 편협한 시각 에서 벗어나 문제를 폭넓게 보아야겠군.
- ③ 문제의 해결을 위해 서로 다른 관점을 비판적으로 통합하여 문제에 대한 생각을 새롭게 구성할 수 있어야겠군.
- ④ 정보를 이해하는 수준을 넘어, 각 관점의 타당성을 검토하고 평가 내용을 통합함으로써 문제를 깊이 이해해야겠군.
- ⑤ 문제에 대한 여러 관점을 다각도로 검토하고, 비판적 판단을 유보함으로써 자신의 관점이 지닌 타당성을 견고히 해야겠군.
- 3. 다음은 윗글을 읽은 학생의 독서 기록장 일부이다. 이에 대한 설명으로 가장 적절한 것은?

나의 독서 대부분은 정보 습득을 위한 것이었다. 책의 내용이 그대로 내 머릿속으로 옮겨져 지식이 쌓이기만을 바랐지 내면의 성장을 생각하지 못했다. 윤동주 평전을 임으며 스스로에게 질문을 던지는 이 시간이 나에 대해 사색하며 삶을 가꾸는 소중한 시간 임을 새삼 느낀다. 오늘 나는 책장을 천천히 넘기며 나에게로의

- ① 삶을 성찰하게 하는 독서의 가치를 깨닫고 이를 실천하려는 모습을 보이고 있다.
- ② 문학 분야에 편중되었던 독서 습관을 버리고 다양한 분야의 책을 읽으려는 노력을 보이고 있다.
- ③ 독서를 지속적으로 실천하지 못한 태도를 반성하고 문제 해결을 위해 장기적인 독서 계획을 세우고 있다.
- ④ 내면적 성장을 위한 도구로서의 독서의 중요성을 인식하고 다양한 매체를 활용한 독서의 방법을 제안하고 있다.
- ⑤ 개인의 지적 성장에 머무는 독서의 한계를 지적하고 타인과 경험을 공유하는 독서 토론의 필요성을 강조하고 있다.

13. 윗글을 참고할 때, <보기>에 대한 반응으로 가장 적절한 **것은? [3점]**

홀수형

--<보 기>---

브레턴우즈 체제가 붕괴된 이후 두 차례의 석유 가격 급등을 겪으면서 기축 통화국인 A국의 금리는 인상되었고 통화 공급은 감소했다. 여기에 A국 정부의 소득세 감면과 군비 증대는 A국의 금리를 인상시켰으며, 높은 금리로 인해 대량으로 외국 자본이 유입되었다. A국은 이로 인한 상황을 해소하기 위한 국제적 합의를 주도하여, 서로 교역을 하며 각각 다른 통화를 사용하는 세 국가 A, B, C는 외환 시장에 대한 개입을 합의 했다. 이로 인해 A국 통화에 대한 B국 통화와 C국 통화의 환율은 각각 50%, 30% 하락했다.

- ① A국의 금리 인상과 통화 공급 감소로 인해 A국 통화의 신뢰도가 낮아진 것은 외국 자본이 대량으로 유입되었기 때문이겠군.
- ② 국제적 합의로 인한 A국 통화에 대한 B국 통화의 환율 하락 으로 국제 유동성 공급량이 증가하여 A국 통화의 가치가 상승
- ③ 다른 모든 조건이 변하지 않았다면, 국제적 합의로 인해 A국 통화에 대한 B국 통화의 환율과 B국 통화에 대한 C국 통화의 환율은 모두 하락했겠군.
- ④ 다른 모든 조건이 변하지 않았다면, 국제적 합의로 인해 A국 통화에 대한 B국과 C국 통화의 환율이 하락하여, B국에 대한 C국의 경상 수지는 개선되었겠군.
- ⑤ 다른 모든 조건이 변하지 않았다면, A국의 소득세 감면과 군비 증대로 A국의 경상 수지가 악화되며, 그 완화 방안 중 하나는 A국 통화에 대한 B국 통화의 환율을 상승시키는 것이겠군.

[14~17] 다음 글을 읽고 물음에 답하시오.

국어 영역

주차하거나 좁은 길을 지날 때 운전자를 돕는 장치들이 있다. 이 중 차량 전후좌우에 장착된 카메라로 촬영한 영상을 이용하여 차량 주위 360°의 상황을 위에서 내려다본 것 같은 영상을 만들어 차 안의 모니터를 통해 운전자에게 제공하는 장치 가 있다. 운전자에게 제공되는 영상이 어떻게 만들어지는지 알아

5

먼저 차량 주위 바닥에 바둑판 모양의 격자판을 펴 놓고 카메라로 촬영한다. 이 장치에서 사용하는 광각 카메라는 큰 시야각을 갖고 있어 사각지대가 줄지만 빛이 렌즈를 @ 지날 때 렌즈 고유의 곡률로 인해 영상이 중심부는 볼록하고 중심부 에서 멀수록 더 휘어지는 현상, 즉 렌즈에 의한 상의 왜곡이 발생한다. 이 왜곡에 영향을 주는 카메라 자체의 특징을 내부 변수라고 하며 왜곡 계수로 나타낸다. 이를 알 수 있다면 왜곡 모델을 설정하여 왜곡을 보정할 수 있다. 한편 차량에 장착된 카메라의 기울어짐 등으로 인해 발생하는 왜곡의 원인을 외부 변수라고 한다. ①촬영된 영상과 실세계 격자판을 비교하면 영상에서 격자판이 회전한 각도나 격자판의 위치 변화를 통해 카메라의 기울어진 각도 등을 알 수 있으므로 왜곡을 보정할 수 있다.

왜곡 보정이 끝나면 영상의 점들에 대응하는 3차원 실세계의 점들을 추정하여 이로부터 원근 효과가 제거된 영상을 얻는 시점 변환이 필요하다. 카메라가 3차원 실세계를 2차원 영상으로 투영하면 크기가 동일한 물체라도 카메라로부터 멀리 있을수록 더 작게 나타나는데, 위에서 내려다보는 시점의 영상에서는 거리에 따른 물체의 크기 변화가 없어야 하기 때문이다.

© 왜곡이 보정된 영상에서의 몇 개의 점과 그에 대응하는 실세계 격자판의 점들의 위치를 알고 있다면, 영상의 모든 점들과 격자판의 점들 간의 대응 관계를 가상의 좌표계를 이용하여 기술할 수 있다. 이 대응 관계를 이용해서 영상의 점들을 격자의 모양과 격자 간의 상대적인 크기가 실세계에서와 동일하게 유지 되도록 한 평면에 놓으면 2차원 영상으로 나타난다. 이때 얻은 영상이 ⓒ위에서 내려다보는 시점의 영상이 된다. 이와 같은 방법으로 구한 각 방향의 영상을 합성하면 차량 주위를 위에서 내려다본 것 같은 영상이 만들어진다.

14. 윗글의 내용과 일치하는 것은?

- ① 차량 주위를 위에서 내려다본 것 같은 영상은 360°를 촬영 하는 카메라 하나를 이용하여 만들어진다.
- ② 외부 변수로 인한 왜곡은 카메라 자체의 특징을 알 수 있으면 쉽게 해결할 수 있다.
- ③ 차량의 전후좌우 카메라에서 촬영된 영상을 하나의 영상으로 합성한 후 왜곡을 보정한다.
- ④ 영상이 중심부로부터 멀수록 크게 휘는 것은 왜곡 모델을 설정하여 보정할 수 있다.
- ⑤ 위에서 내려다보는 시점의 영상에 있는 점들은 카메라 시점의 영상과는 달리 3차원 좌표로 표시된다.

국어 영역





- - 한국교육과정평가원이 출제하는 '국어 영역'은

문해 능숙도 혹은 언어적인 사고력을 측정하는 한국어판 문해 평가이다.



국어 영역





AI가 풀어본다면

한국교육과정평가원이 출제하는 '국어 영역'은

문해 능숙도 혹은 언어적인 사고력을 측정하는 <u>한국어판 문해 평가</u>이다.

도로보군



인공지능 '도로보군'은 왜 도전을 멈추었는가?

#512 | 권재원, 실천교육교사모임 고문 | 2020.01.22

日 AI 도로보군, 200점 만점 대입서 185점...내 년 도쿄대 입학 노려

명문대에 합격하는 인공지능

2016년, 일본은 물론 세계를 떠들썩하게 만든 사건이 일어났다. 인공: 의 대학입학 시험인 센터 시험에서 상위 20%에 해당되는 성적을 거둔 험을 응시조차 하지 않는 것을 감안하면 사실상 상위 10%, 즉 우리나라 \exists 로벌인재포럼 2020 학'에 능히 입학할 정도의 성적이다.

더구나 이 인공지능 로봇은 눈과 손을 사용했다. 렌즈를 통해 인쇄된 시험 뒤 문제를 풀고, 그 결과를 펜이 장착된 로봇 팔을 이용하여 답안지에 직접 2016년 95점 비해 두배 **늘어** 해당 숫자를 출력하는 것이 아니라 시험지를 읽고 해당되는 답 칸에 직접 ㄸ 술식 문항이 출제된다) 직접 풀이과정을 답지에 썼다.

도로보군 이전에 사람과 대결하여 센세이션을 일으켰던 왓슨 그리고 알파고! 지 알 수 있다. 왓슨은 퀴즈쇼 저파디에서 사람을 물리치고 우승을 차지했고, 음성을 합성하여 답하는 방식이 아니라 키보드를 통해 문제를 입력 받으면 답을 했다. 즉 신체 없이 지능으로만 작성했다. 알파고 역시 렌즈를 통해 바둑 돌의 우 위에 돌을 놓는 대신 사람이 마우스로 상대방의 착점을 찍어 입력해주면 다음 최 작동했다. 실제 바둑판에 그 돌을 놓으려면 역시 사람의 도움이 필요했다.

의 학생들을 문자 그대로 '로봇 만도 못한' 녀석들로 만들어 버렸다. 그럼에도 불구하 주인공이다. 도쿄대의 '도'와 '로봇', 어린 남자의 호칭인 '군'을 합쳐서 지은 도달하지 못했다. 도로보군 프로젝트이 의권 모르 도달하지 못했다. 도로보군 프로젝트의 원래 목표는 도쿄대에 합격하는 것이었다. 출 50% 수준에도 미치지 못했던 것과 비교하면 장족의 발전이다.

정영효기자 🗘

입력 2020.11.09 17:28 수정 2020.11.20 15:45 지면 A10

ナナ

AI 경쟁력 회복 속도내는 일본

몇초면 문장 전체 맥락 파악 딥러닝후 문장 독해력 크게 늘어



'인공지능(AI)이 일본 최고 명문 도쿄대의 입학시험을 치른다면?'

구글이 개발한 AI 바둑 프로그램인 알파고가 이세돌 9단을 4 대 1로 꺾으면서 반면 도로보군은 사람과 똑같은 조건에서 시험을 쳤다. 그것도 바둑보다 훨씬 복집 인류에게 충격을 준 2016년 3월, 일본에서는 Al를 도쿄대에 입학시키는 시험을. 사람의 도움 없이 시험지를 직접 이어고 기업으로 모르는 이를 모르고 있다. 그런 그리저너하여구소(NII)가 개발한 Al 도로보 프로젝트가 한창이었다. 일본 국립정보학연구소(NII)가 개발한 AI 도로보군이 이름이다. 2011년 시작해 연구자만 100명 넘게 투입된 이 프로젝트는 2021년까지 도로보군을 도쿄대에 합격시키는 것이 목표다.

오늘의 주요뉴스



단독 "월세만 찾아요"...외국인



노벨경제학상에 버냉키 등 3 명..."금융위기 대처능력 높...



"일본보다 나쁘다"...한국, 6개 월째 경기하강 '초비상'



하워드 막스 "지금 시장 떠나 는 건 실수" 한경 글로벌마...



푸틴 최측근 "크림대교 보복은



이 시각 관심정보 ^{AD}

한국로또 조작 밝혀졌다!? 용지 속 "QR코드" 찍... 역류성식도염, 두통있다면 검사 필수! 간단치료... 역류성식도염 증상 "부작용없는 치료법" 개발!?

- 1. 윗글을 바탕으로 할 때, ①의 답으로 적절하지 않은 것은?
- 2. <보기>는 ②과 같이 독서하기 위해 학생이 찾은 독서 방법이다. 이에 대한 반응으로 적절하지 않은 것은? [3점]
- 3. 다음은 윗글을 읽은 학생의 독서 기록장 일부이다. 이에 대한 설명으로 가장 적절한 것은?
- 4. (가)와 (나)에 대한 설명으로 가장 적절한 것은?
- 5. (가)에서 알 수 있는 헤겔의 생각으로 적절하지 않은 것은?
- 6. (가)에 따라 직관·표상·사유 의 개념을 적용한 것으로 적절 하지 않은 것은?
- 7. (나)의 글쓴이의 관점에서 ①과 ①에 대한 헤겔의 이론을 분석한 것으로 적절하지 않은 것은?
- 8. <보기>는 헤겔과 (나)의 글쓴이가 나누는 가상의 대화의 일부 이다. ②에 들어갈 내용으로 가장 적절한 것은? [3점]
- 9. 문맥상 ⓐ~@와 바꾸어 쓰기에 가장 적절한 것은?
- 10. 윗글을 통해 답을 찾을 수 없는 질문은?
- 11. 윗글을 바탕으로 추론한 내용으로 적절하지 않은 것은?
- 12. 미국을 포함한 세 국가가 존재하고 각각 다른 통화를 사용할 때, ¬~□에 대한 설명으로 적절한 것은?
- 13. 윗글을 참고할 때, <보기>에 대한 반응으로 가장 적절한 것은? [3점]

- 6. (가)에 따라 직관·표상·사유 의 개념을 적용한 것으로 적절 하지 않은 것은?
 - ① 먼 타향에서 밤하늘의 별들을 바라보는 것은 직관을 통해, 같은 곳에서 고향의 하늘을 상기하는 것은 표상을 통해 이루어 지겠군.
 - ② 타임머신을 타고 미래로 가는 자신의 모습을 상상하는 것과, 그 후 판타지 영화의 장면을 떠올려 보는 것은 모두 표상을 통해 이루어지겠군.
 - ③ 초현실적 세계가 묘사된 그림을 보는 것은 직관을 통해, 그 작품을 상상력 개념에 의거한 이론에 따라 분석하는 것은 사유를 통해 이루어지겠군.
 - ④ 예술의 새로운 개념을 설정하는 것은 사유를 통해, 이를 바탕으로 새로운 감각을 일깨우는 작품의 창작을 기획하는 것은 직관을 통해 이루어지겠군.
 - ⑤ 도덕적 배려의 대상을 생물학적 상이성 개념에 따라 규정 하는 것과, 이에 맞서 감수성 소유 여부를 새로운 기준으로 제시하는 것은 모두 사유를 통해 이루어지겠군.





(3)

This is a picture of two teddy bears on the moon.

What are they doing?

- They are having a conversation.
 - What object are they using?
- It looks like a computer.
 - Is this surprising?
- Yes, it is surprising.

Why is this picture surprising to you?

I think it is surprising because teddy bears are not usually found on the moon.

- 6. (가)에 따라 직관·표상·사유 의 개념을 적용한 것으로 적절 하지 않은 것은?
 - ① 먼 타향에서 밤하늘의 별들을 바라보는 것은 직관을 통해. 같은 곳에서 고향의 하늘을 상기하는 것은 표상을 통해 이루어 지겠군.
 - ② 타임머신을 타고 미래로 가는 자신의 모습을 상상하는 것과, 그 후 판타지 영화의 장면을 떠올려 보는 것은 모두 표상을 통해 이루어지겠군.
 - ③ 초현실적 세계가 묘사된 그림을 보는 것은 직관을 통해. 그 작품을 상상력 개념에 의거한 이론에 따라 분석하는 것은 사유를 통해 이루어지겠군.
 - ④ 예술의 새로운 개념을 설정하는 것은 사유를 통해, 이를 바탕 으로 새로운 감각을 일깨우는 작품의 창작을 기획하는 것은 직관을 통해 이루어지겠군.
 - ⑤ 도덕적 배려의 대상을 생물학적 상이성 개념에 따라 규정 하는 것과, 이에 맞서 감수성 소유 여부를 새로운 기준으로 제시하는 것은 모두 사유를 통해 이루어지겠군.

어쩌면...?

NATURAL LANGUAGE PROCESSING

Use tokenizers from



Tokenizers

Inference for multilingual models

TASK GUIDES

Text classification

Token classification

Question ar wering

Language modeling

Translation

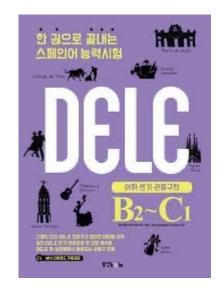
Summarization

Multiple choice

- 6. (가)에 따라 직관·표상·사유 의 개념을 적용한 것으로 적절 하지 않은 것은?
 - ① 먼 타향에서 밤하늘의 별들을 바라보는 것은 직관을 통해. 같은 곳에서 고향의 하늘을 상기하는 것은 표상을 통해 이루어 지겠군.
 - ② 타임머신을 타고 미래로 가는 자신의 모습을 상상하는 것과, 그 후 판타지 영화의 장면을 떠올려 보는 것은 모두 표상을 통해 이루어지겠군.
 - ③ 초현실적 세계가 묘사된 그림을 보는 것은 직관을 통해. 그 작품을 상상력 개념에 의거한 이론에 따라 분석하는 것은 사유를 통해 이루어지겠군.
 - ④ 예술의 새로운 개념을 설정하는 것은 사유를 통해, 이를 바탕 으로 새로운 감각을 일깨우는 작품의 창작을 기획하는 것은 직관을 통해 이루어지겠군.
 - ⑤ 도덕적 배려의 대상을 생물학적 상이성 개념에 따라 규정 하는 것과, 이에 맞서 감수성 소유 여부를 새로운 기준으로 제시하는 것은 모두 사유를 통해 이루어지겠군.





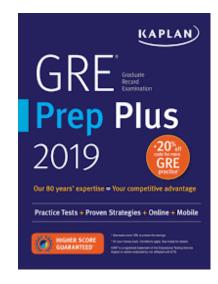












고려사항

- 데이터는 충분한가 (대부분의 문제의 경우 hwp(한글파일) 존재)
 (6월,9월,11월(수능)만 따질 경우16년x3개월x45문제 = 2160문항 / 부족하다면 평가염, 고1, 고2, LEET)
- · 그래서 그 입력에 해당하는 문제는 잘 푸나? (BERT for Multiple Choice)
- · 사진이나 도표는 어떻게 할 것인가 (VQA)
- · 토큰의 수가 4096개를 넘어가는 long text의 경우 어떻게 처리할 것인가?
- · 질문의 유형 혹은 문학/ 비문학/ 독서와 문법 등으로 나눠서 해결 할 것인가?
- · 고전시가의 옛말, 한자어 등의 경우는 어떻게 처리할 것인가?

- ♦ "고3"을 위한 공간
 - 3학년 모의고사 전과목 자료
 - 07~22년 국영수(평가원)
 - 07~22년 국영수(교육청)
 - 07~22년 과학탐구(평가원)
 - 07~22년 사회탐구(평가원)
 - 07~22년 사탐, 과탐(교육청)
 - 2004~2006년 수능/모의고사
- ◆ "고2"를 위한 공간
 - 2학년 모의고사 전과목 자료
 - 08~22년 국영수 문제
 - 08~22년 사탐,과탐 문제
 - 2004~2007년 모의고사
- ♦ "고1"을 위한 공간
 - 1학년 모의고사 전과목 자료
 - 08~22년 국영수 문제
 - 08~22년 사탐,과탐 문제
 - 2004~2007년 모의고사

- Multimodal
- Text Summarization
- QA
- Language Generation
- Chat Bot
- MLOps / End-to-End Service

다양한 분야에 관심 많습니다!!

재밌고, 알차게, 끝까지 프로젝트를 진행해보고 싶습니다.

감사합니다