

# 디코더를 활용한 기계독해 모델의 근거 추출 방법

한규빈<sup>01</sup>, 장영진<sup>2</sup>, 김학수<sup>2</sup>

건국대학교 전기전자공학부<sup>1</sup>, 건국대학교 인공지능학과<sup>2</sup>  
{rbqlsqf, danyon, nlprkim}@konkuk.ac.kr

## Evidence Extraction Method for Machine Reading Comprehension Model using Recursive Neural Network Decoder

Kyubeen Han <sup>01</sup>, Youngjin Jang, Harksoo Kim  
Department of Electrical and Electronics Engineering, Konkuk University<sup>1</sup>,  
Department of Artificial Intelligence, Konkuk University<sup>2</sup>

### 요약

최근 인공지능 시스템이 발전함에 따라 사람보다 높은 성능을 보이고 있다. 또한 전문 지식에 특화된 분야 (질병 진단, 법률, 교육 등)에도 적용되고 있지만 이러한 전문 지식 분야는 정확한 판단이 중요하다. 이로 인해 인공지능 모델의 결정에 대한 근거나 해석의 중요성이 대두되었다. 이를 위해 설명 가능한 인공지능 연구인 XAI가 발전하게 되었다. 이에 착안해 본 논문에서는 기계독해 프레임워크에 순환 신경망 디코더를 활용하여 정답 뿐만 아니라 예측에 대한 근거를 추출하고자 한다. 실험 결과, 모델의 예측 답변이 근거 문장 내 등장하는지에 대한 실험과 분석을 수행하였다. 이를 통해 모델이 추론 과정에서 예측 근거 문장을 기반으로 정답을 추론한다는 것을 확인할 수 있었다.

주제어: XAI, 인공지능, 기계독해, 근거 추출

### 1. 서론

XAI(explainable Artificial Intelligence)란 사용자에게 인공지능 시스템의 결과를 이해시키거나 사용자의 의사 결정을 지원하기 위해 설명을 제공하는 연구분야이다. XAI는 2004년 군사 시뮬레이션 및 컴퓨터 게임의 인공지능 시스템 결과를 분석하기 위해 처음 도입되었으며[1], 인공지능 시스템의 동작을 이해하고 설명할 수 있는 방법을 연구한다[2]. 사용자는 XAI를 통해 모델의 결정 방식을 이해함으로써 예측 과정에 대한 신뢰성을 제공할 수 있다[3]. 또한 인공지능이 인간보다 뛰어날 경우, 인간은 인공지능의 예측에 대한 해석을 통해 새로운 지식을 배울 수 있고, 숨겨진 법칙을 발견할 수 있다[4].

최근 인공지능 시스템이 급격하게 발전하면서 사람보다 뛰어난 시스템이 공개되었다[5-7]. 하지만 높은 성능 향상에도 불구하고 실제 모델의 예측 결과를 신뢰할 수 있는지에 대한 지적이 발생했으며[8-10], 이는 전문 지식(질병 진단, 법률 등)과 관련된 시스템에서 더욱 중요한 문제로 대두되었다[11-12]. 하지만 인공지능 시스템의 특성상 모델 내부의 파라미터를 직접 해석하는데 많은 어려움[10]이 있기 때문에 최근에는 XAI에 대한 연구가 활발히 진행되고 있다.

본 논문에서는 순환 신경망 디코더(Recurrent Neural Network; RNN)[13]를 활용하여 시스템의 예측에 대한 설명을 제공할 수 있는 연구를 수행하고자 하며, 이를 높은 수준의 자연어 이해 능력이 필요한 기계독해(Machine Reading Comprehension; MRC) 작업에 적용하고자 한다.

### 2. 관련 연구

초기 자연어 처리 분야의 연구는 의사 결정 트리, 히든 마르코프 모델, 로지스틱 회귀 등과 같이 동작 원리를 이해하기 쉬운 화이트박스 시스템 기반으로 발전되었다[14]. 최근에는 시스템 내부 파라미터를 해석하기 어렵지만 높은 성능을 보이는 블랙박스 시스템 기반의 딥러닝 모델을 주축으로 연구가 진행되고 있다. 이러한 블랙박스 시스템은 전문 지식에 특화된 분야에도 활용되고 있다[15-19]. 하지만 이러한 전문 지식 분야에서는 정확한 판단이 중요하기 때문에 인공지능 모델의 결정에 대한 근거나 해석이 필요하다[2]. 위와 같이 모델의 신뢰성과 해석 관점에서 XAI는 매우 중요해졌으며, 이에 대한 다양한 연구가 수행되었다. [20]는 자연어 추론 작업에서 전제 문장과 가설 문장 사이의 대응관계(alignment)를 분석했으며, 모델의 학습을 완료한 후, 학습 가능한 마스크와 KLDivLoss(Kullback-Leibler divergence loss) 함수를 통해 토큰 단위의 예측 근거를 제공해준다. [21]은 멀티홉(multi-hop) 질의에 대한 응답을 찾기 위한 근거를 토큰 단위가 아닌 문장 단위로 추출한다. 문장 단위를 통해 추출된 근거는 추론 과정을 더 명시적으로 나타낼 수 있고 이를 통해 모델이 실제 답을 어떻게 도출하는지 알 수 있다.

본 논문에서는 모델의 학습을 완료한 후 마스크를 추가 학습하는 [20]과 달리 한 번의 학습과정으로 예측에 대한 근거를 추출하고자 한다. 또한 [21]에서 추론 과정을 명시적으로 나타내기 위해 문장 단위의 근거를 추출한 점에서 착안하여 문장 단위의 근거를 추출한다.

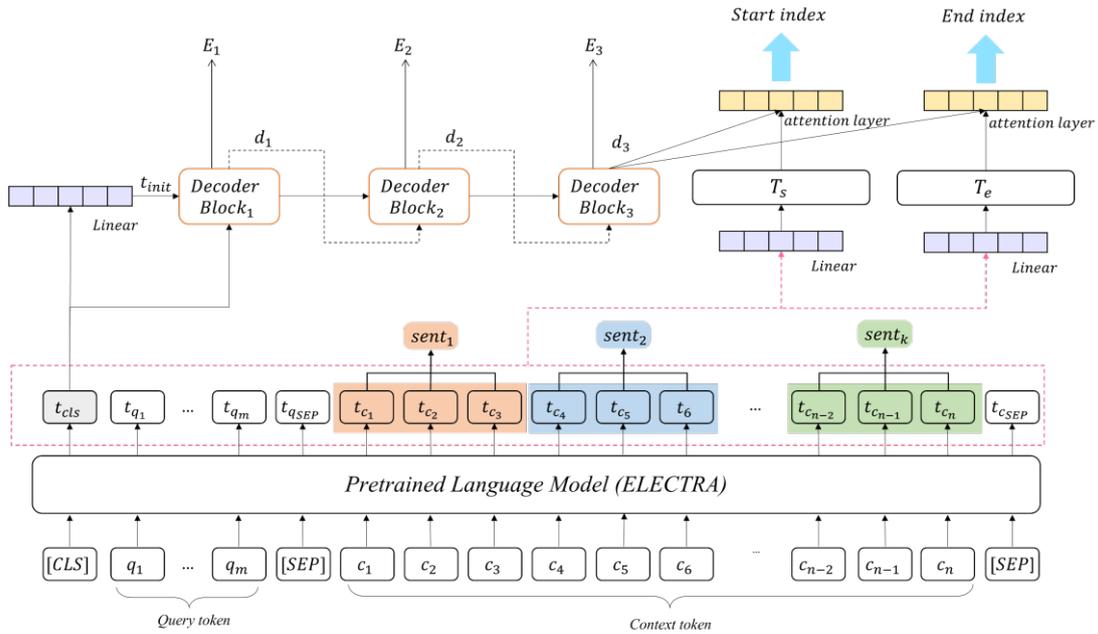


그림 1. 제안 모델 구조도

### 3. 제안 모델

제안 모델 구조는 그림 1과 같다. 제안 모델은 기존 기계독해 프레임워크[22]에 예측에 대한 근거를 제공하기 위한 Decoder Block을 추가하였다. Decoder Block의 구조는 그림 2와 같다. Decoder Block은 GRU[23] 기반의 포인터 네트워크[24]를 통해 추론에 필요한 정보를 축적한다. 제안 모델은 Decoder Block를 통해 추출된 문장 정보를 제공함으로써, 사용자에게 예측에 대한 근거를 보여준다. Decoder Block의 길이는 3으로 제한하며, 사용자에게 고정 개수의 근거 문장을 제공한다.

#### 3.1 모델 입력

제안 모델의 입력은 기존 기계독해 프레임워크와 동일하게 토큰화된 질의와 문단을 [CLS]기호와 [SEP]기호로 구분하여  $\{[CLS], q_1, \dots, q_m, [SEP], c_1, c_2, \dots, c_n, [SEP]\}$  형태로 구성한다. 이때,  $m$ 과  $n$ 은 각각 토큰화된 질의와 문단의 토큰 수이다. 이후, 사전 학습 언어모델을 통해 문맥 정보가 반영된 토큰 벡터 열  $T = \{t_{cls}, t_{q_1}, \dots, t_{q_m}, \dots, t_{c_n}, t_{c_{SEP}}\}$ 를 생성한다. 3절에서 언급한대로 예측에 대한 근거를 문장 단위로 추출하기 위하여 토큰 벡터 열  $T$ 로부터 문장 벡터를 생성한다. 구체적으로, 같은 문장에 해당하는 토큰 벡터에 대해 mean-pooling을 계산하여 문장 벡터  $S = \{sent_1, sent_2, \dots, sent_k\}$ 를 생성한다. 이때  $k$ 는 모델 입력의 문단을 구성하는 문장 개수이다.

#### 3.2 Decoder

본 논문에서 Decoder Block의 초기 입력과 초기 상태 값은 각각  $t_{cls}$ 과  $t_{init}$ 으로 설정했다. 이때  $t_{init}$ 은 아래의 수식(1)과 같이 계산된다.

$$t_{init} = W_1 t_{cls} \quad (1)$$

위 수식(1)의  $W_1$ 은 학습 가능한 가중치 행렬을 의미한다.

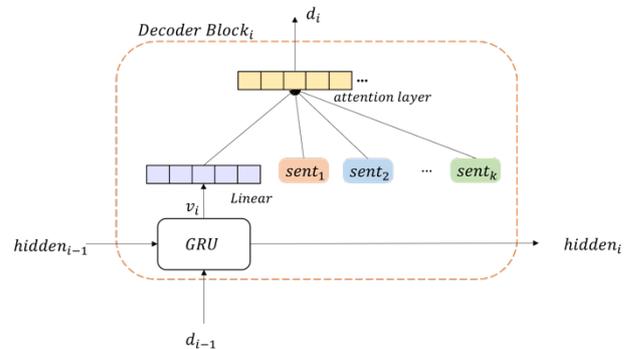


그림 2. Decoder Block 구조

Decoder Block의 내부 구조는 그림 2와 같으며, GRU의 출력 벡터  $v_i$ 는 아래의 수식(2)를 통해 계산된다.

$$v_i = GRU(d_{i-1}, hidden_{i-1}) \quad (2)$$

$i$ 는 디코딩 step 수를 의미하며,  $d_{i-1}$ 와  $hidden_{i-1}$ 는 각각 이전 time step의 Decoder Block 출력 값과 GRU 상태 값을 의미한다. GRU를 통해 생성된 출력 값  $v_i$ 는 문장 벡터  $S$ 와의 Attention을 계산하여 정답 추론에 필요한 문장정보를 반영하여  $d_i$ 를 생성한다. 이에 대한 수식은 아래의 (3)과 같다.

$$score(S, v_i) = S^T W_2 v_i$$

$$\alpha_i = \frac{\exp(score(S, v_i))}{\sum_{i'=1}^i \exp(score(S, v_{i'}))} \quad (3)$$

$$d_i = \sum_s \alpha_i S$$

위의 수식에서  $W_2$ 는 학습 가능한 가중치 행렬을 의미한다. 정답 추론 학습하는 과정은 다음과 같다. 사전 학습 언어 모델의 출력 벡터 열  $T$ 를 학습 가능한 가중치 행렬  $W_3$ 와  $W_4$ 를 사용하여  $T_s$ 와  $T_e$ 를 만든다.

$$T_s = W_3(T) \in \mathbb{R}^{L \times H} \quad (4)$$

$$T_e = W_4(T) \in \mathbb{R}^{L \times H}$$

위 수식에서  $L$ 은 모델의 입력 길이를 의미하고  $H$ 는 단어 벡터의 차원 수를 의미한다. 위의 수식 (4)를 통해 생성된  $T_s$ 와  $T_e$ 는 Decoder Block의 마지막 출력 값  $d_3$  사이의 Attention 연산을 통해 정답 위치를 학습한다.

$$p_s = T_s^T d_3 \in \mathbb{R}^L \quad (5)$$

$$p_e = T_e^T d_3 \in \mathbb{R}^L$$

위 수식에서  $p_s$ 와  $p_e$ 는 모델의 입력 토큰에 대한 정답 위치 시작, 끝 점에 대한 점수를 의미한다. 제안 모델은 Cross Entropy 함수를 통해 학습된다. 이에 대한 수식은 아래의 식 (6)과 같다.

$$L_{CE}^s = - \sum_{i=1}^n y_s \log(p_s)$$

$$L_{CE}^e = - \sum_{i=1}^n y_e \log(p_e) \quad (6)$$

$$L_{total} = L_{CE}^s + L_{CE}^e$$

학습된 제안모델은 추론 단계에서 아래의 수식 (7)을 통해 근거 문장을 추출한다.

$$E_i = \operatorname{argmax}(\alpha_i) \quad (7)$$

## 4. 실험 및 결과

### 4.1 실험 준비

실험에 사용한 데이터는 KorQuAD(Korean Question Answering Dataset) v1.0[25]이다. 실험 데이터는 여러 문장으로 구성된 문서, 질의, 정답의 시작 인덱스, 정답 어휘로 구성되어 있으며, 학습 데이터 60,407개, 평가

데이터 5,774개의 질의응답쌍으로 구분되어 있다. 본 논문에서는 KoELECTRA-Small-v2-Discriminator[26]를 인코더로 사용하였다. 본 실험에서는 성능 지표로 Exact Match와 F1-score를 사용했다. Exact Match는 모델이 예측한 답변과 정답 사이의 완전 일치도를 의미하며, F1-score는 모델이 예측한 답변과 정답 사이의 부분 일치도를 의미한다.

### 4.2 실험 결과 및 분석

본 논문에서 제안하는 Decoder Block이 정답 추론에 어떤 영향을 끼치는지 확인하기 위해 질의에 대한 정답 추론 성능을 기존 기계독해 프레임워크와 비교하였다. 이에 대한 실험 결과는 아래 표 1과 같다.

표 1. 정답 추론 성능표

	Exact Match (%)	F1-score (%)
Ours	81.61	90.55
Baseline	81.35	90.40

위 실험 결과에 따르면 본 논문에서 제안한 Decoder Block이 정답 추론에 기여하는 것을 확인할 수 있었다. 다음으로 Decoder Block에서 추출된 근거 문장이 정답 추론에 어떤 기여를 하는지 분석하기 위해 아래의 실험을 진행했다. 우선 본 논문은 정답을 포함하는 단일 문장을 근거 문장(Silver Evidence Sentence)이라고 가정한다. 위 가정이 타당한지 확인하기 위해서 우리는 Silver Evidence Sentence만으로 정답 추론이 가능한지 확인하였다. 이에 대한 실험 결과는 표 2와 같다.

표 2. Silver Evidence Sentence 내 정답 추론 성능표

	Exact Match (%)	F1-score (%)
Baseline	83.77	92.43

표 1의 Baseline F1-score와 표 2의 F1-score를 비교했을 때 표 2의 결과가 더 높다는 것을 알 수 있다. 이를 통해 Silver Evidence Sentence만으로도 정답 추론이 가능하다고 말할 수 있으며, 이는 Silber Evidence Sentence가 정답 추론에 필요한 근거 문장으로 볼 수 있다. 이 가정을 바탕으로 우리는 Decoder Block에서 추출된 3개의 근거 문장  $E = \{E_1, E_2, E_3\}$ 를 가지고 정답 추출 실험을 진행했다. 이에 대한 실험 결과는 아래와 같다.

표 3. 예측 근거 문장  $E$  내 정답 추론 성능표

	Exact Match (%)	F1-score (%)
Ours(evidence)	77.00	86.20

표 3의 결과는 근거 문장  $E$ 만을 통해 정답을 추론할

때 성능이 표 2의 결과보다 떨어지는 것을 확인할 수 있는데 이는 Silver Evidence Sentence와 비교하여 예측 근거 문장 **E**에 실제 정답이 비교적 덜 포함하고 있을 것이라 판단했다. 이를 보다 정확하게 분석하기 위해 예측 근거 문장 **E** 내부에 정답 어휘가 등장하는지 확인했다. 이에 대한 결과는 표 4와 같다.

표 4. 예측 근거 문장 **E** 내 정답 어휘 Recall@N 성능표

	Recall@1	Recall@2	Recall@3
Ours	91.08	91.63	91.92

위 표 4에서 사용한 평가 지표인 Recall@N은 N개의 문장 내에 정답이 포함되는지 측정된 점수이다. 해당 실험을 통해 예측 근거 문장 **E** 내부에 실제 정답이 약 8%정도 포함되지 않는 것을 확인할 수 있었다. 아래의 표 5는 예측 근거 문장 **E**에 대한 실제 예시이다.

표 5. 모델 실제 출력 예시 1

질의	포켓몬스터 금은이 일본에서 발매된 해는?
Silver Evidence Sentence	<b>1999년</b> 일본에서 처음 발매되었으며, 2000년 오스트레일리아와 북아메리카, 2001년 유럽, 2002년 대한민국에서 발매하였다.
예측 근거 문장 <b>E</b>	닌텐도는 《금·은》을 닌텐도 DS용으로 리메이크한 《포켓몬스터 하트 골드·소울 실버》를 <b>2009년</b> 9월 12일 일본에서 발매하였다.

표 5에서 질의에 대한 실제 정답은 Silver Evidence Sentence의 ‘1999년’ 이지만, 모델이 실제로 예측한 정답은 예측 근거 문장 **E**의 ‘2009년’을 예측하였다. 모델이 정답을 틀리는 경우에도 예측 근거 문장 **E**에서 정답을 추론하는 것을 확인할 수 있다. 표 5와 같이 실제 정답과 관계없이 예측 정답에 대한 예측 근거 문장이 타당한지를 확인하기 위해 아래 실험을 진행하였다. 아래 실험은 예측 근거 문장 **E** 내부에 예측 어휘가 등장하는지 확인했다.

표 6. 예측 근거 문장 **E** 내 예측 어휘 Recall@N 성능표

	Recall@1	Recall@2	Recall@3
Ours	91.80	92.34	92.64

표 4와 표 6을 비교했을 때 표 6의 성능이 더 높은 것으로 보아, 실제로 모델의 예측 답변이 실제 정답과 다를 때, 근거 문장을 기반으로 추출한다는 것을 확인할 수 있다.

표 7. 모델 실제 출력 예시 2

질의	1976년 현대산업개발에서 건설하고 현재 국내에서 가장 비싼 아파트 가운데 하나인 아파트의 이름은 무엇인가?
Silver Evidence Sentence	<b>압구정 현대아파트</b> 는 1976년 현대산업개발에서 건설한 아파트단지이다.
예측 근거 문장 <b>E</b>	고소득층 아파트의 상징이 된 <b>압구정동 현대아파트</b> 는 30여 년이 흐른 지금도 국내에서 가장 비싼 아파트 가운데 하나다.

표 8. 모델 실제 출력 예시 3

질의	바르셀로나 올림픽에서 소토마요르가 2.39m를 통과한 라운드는?
Silver Evidence Sentence	<b>마지막 라운드</b> 에서 도약하면서, 그는 2.37m에서 첫 2개의 시도들에서 실패하였다.
예측 근거 문장 <b>E</b>	나머지의 모든 선수들이 자신의 3번째 시도들을 놓칠 때에 소토마요르는 금메달에 보장하여 3번째 시도를 통과하여 자신이 놓친 2.39m의 기록 높이에서 하나의 시도를 위하여 선택하였다.
예측 답변 어휘 포함 문장	<b>동점 결승전</b> 의 근거에서 소토마요르가 우승 선수로 선언되었으며, 세베리는 은메달, 폴란드의 아르투르 파르티카, 오스트레일리아의 팀 포사이스와 미국의 홀리스 콘웨이는 동메달을 획득하게 된다.

표 7과 표 8은 질의에 대한 정답을 찾기 위해서 2개 이상의 문장을 참조해야 하는 경우이다. 해당 결과를 토대로 제안 모델의 Decoder Block은 문제 해결을 위해 둘 이상의 문장을 참조하지 않고 하나의 문장을 반복적으로 추출하는 경향이 있음을 알 수 있다. 위 결과를 토대로 추후 연구에서는 예측 근거 문장의 추론 과정에서 중복 추출을 막거나, 정보량에 따른 근거 문장 개수를 유동적으로 조절할 계획이다.

### 5. 결론 및 향후 연구

본 논문은 추론 과정을 명시적으로 나타낼 수 있는 문장 단위의 정보를 추출하여 모델의 학습 과정을 분석하였다. 분석 결과, 근거 문장이 모델의 예측 답변을 포함하는 경우가 92.64%였다. 이를 통해 모델의 예측 과정이 근거 문장을 통해 이루어진다는 것을 확인할 수 있었다. 또한 데이터 분석을 통해 기계적해 모델이 두 가지 이상의 문장을 참조해야 하는 질의에 대해서 다수의 근거 문장을 추출하지 못한다는 한계점을 분석했다. 이는 제안 모델이 Decoder Block에서 중복 추출을 허용했기 때문이라고 판단되며, 또한 정보량에 따른 근거 문장을 유동적으로 조절하지 않았다는 점에서 한계점이 있다.

추후에 Decoder Block에서 추출된 예측 근거 문장 **E** 중에서 더 효과적인 근거 문장을 분석하며, 근거 문장 수를 정보량에 따라 유동적으로 조절할 수 있는 모델을 실험해볼 계획이다.

### 감사의 글

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2022-0-00369, (4세부) 전문지식 대상 판단결과의 이유/근거를 설명가능한 전문가 의사결정 지원 인공지능 기술개발) 본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 「대학혁신지원사업」의 연구결과임

### 참고문헌

- [1] M. van Lent, W. Fisher, and M. Mancuso, "An explainable artificial intelligence system for small-unit tactical behavior," AAAI Conference on Artificial Intelligence, 2004.
- [2] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv: Machine Learning, 2017.
- [3] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (xai)," IEEE Access, Vol. 6, pp. 52 138-52 160, 2018.
- [4] T. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 1412-1421, Sep. 2015.
- [5] OpenAI, "Gpt-4 technical report," 2023.
- [6] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171-4186, Jun. 2019.
- [7] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pp. 2227-2237, Jun. 2018.
- [8] W. J. von Eschenbach, "Transparency and the black box problem: Why we do not trust ai," Philosophy and Technology, Vol. 34, No. 4, pp. 1607-1622, 2021.
- [9] H. Felzmann, E. Fosch-Villaronga, C. Lutz, and A. Tam`o-Larrieux, "Towards transparency by design for artificial intelligence," Science and Engineering Ethics, Vol. 26, No. 6, pp. 3333-3361, 2020.
- [10] M. Abdar, F. Pourpanah, S. Hussain, D. Rezazadegan, L. Liu, M. Ghavamzadeh, P. Fieguth, X. Cao, A. Khosravi, U. R. Acharya, V. Makarenkov, and S. Nahavandi, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," Information Fusion, Vol. 76, pp. 243-297, 2021.
- [11] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," ACM Comput. Surv., Vol. 51, No. 5, aug 2018.
- [12] D. Martens, B. Baesens, T. Van Gestel, and J. Vanthienen, "Comprehensible credit scoring models using rule extraction from support vector machines," Eur. J. Oper. Res., vol. 183, no. 3, pp. 1466-1476, 2007.
- [13] K. Cho, B. van Merri`enboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1724-1734, Oct. 2014.
- [14] O. Loyola-Gonz´alez, "Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view," IEEE Access, Vol. 7, pp. 154096-154113, 2019.
- [15] T. Ward, P. Mascagni, A. Madani, N. Padoy, S. Perretta, and D. Hashimoto, "Surgical data science and artificial intelligence for surgical education," Journal of Surgical Oncology, Vol. 124, pp. 221-230, 08 2021.
- [16] John Murawski. 2019. Mortgage Providers Look to AI to Process Home Loans Faster. Wall Street Journal (18 March 2019). Retrieved 16-September2020 from <https://www.wsj.com/articles/mortgage-providers-look-to-ai-to-process-home-loans-faster-11552899212>
- [17] DJ Pangburn. 2019. Schools are using software to help pick who gets in. what could go wrong? <https://www.fastcompany.com/90342596/schools-are-quietlyturning-to-ai-to-help-pick-who-gets-in-what-could-go-wrong>
- [18] Ben Dattner, Tomas Chamorro-Premuzic, Richard Buchband, and Lucinda Schettler. 2019. The Legal and Ethical Implications of Using AI in Hiring. Harvard Business Review (25 April 2019). Retrieved 26-August-2019 from <https://hbr.org/2019/04/the-legal-and-ethical-implications-of-using-ai-in->

hiring

- [19] Karen Hao. 2019. AI Is Sending People to Jail—And Getting It Wrong. Retrieved February 20, 2021 from <https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/>. Accessed: 2021-02-20.
- [20] Z. Jiang, Y. Zhang, Z. Yang, J. Zhao, and K. Liu, “Alignment rationale for natural language inference,” Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pp. 5372-5387, Aug. 2021.
- [21] J. Chen, S. ting Lin, and G. Durrett, “Multi-hop question answering via reasoning chains,” arXiv preprint, abs/1910.02610, 2021.
- [22] [https://github.com/huggingface/transformers/blob/v4.33.0/src/transformers/models/electra/modeling\\_electra.py#L1341](https://github.com/huggingface/transformers/blob/v4.33.0/src/transformers/models/electra/modeling_electra.py#L1341)
- [23] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, “Empirical evaluation of gated recurrent neural networks on sequence modeling,” In Proceedings of the Neural Information Processing Systems Workshop on Deep Learning. 2014.
- [24] Vinyals, O., M. Fortunato, and N. Jaitly. (2015). “Pointer networks” . In: Advances in Neural Information Processing Systems. 2692-2700
- [25] S. Lim, M. Kim, and J. Lee, “KorQuAD: Korean QA Dataset for Machine Comprehension,” Proc. of the KIISE Korea Software Congress 2018, pp. 539-541, 2018. (in Korean)
- [26] K. Clark, M.-T. Luong, Q. V. Le, and C. D. Manning, “Electra: Pre-training text encoders as discriminators rather than generators,” ICLR, 2020.