

XH-DQN
사실 검증을 위한
그래프 Transformer와 DQN 결합 모델

서민택 나승훈 신동욱 김선훈 강인호
전북대학교 네이버

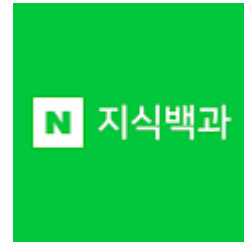
사실 검증 문제

질문 : 2개월 된 강아지는 잠을 얼마나 자?

답변 : 강아지 하루 평균 수면 시간은 15시간 정도 입니다



WIKIPEDIA
La enciclopedia libre



문서 제목 : 반려견의 수면

증거문장 : 강아지는 약 17~18시간 정도 잠을 잡니다



거짓

사실 검증 문제

질문과 답변이 주어짐



문서 유사도

문서검색



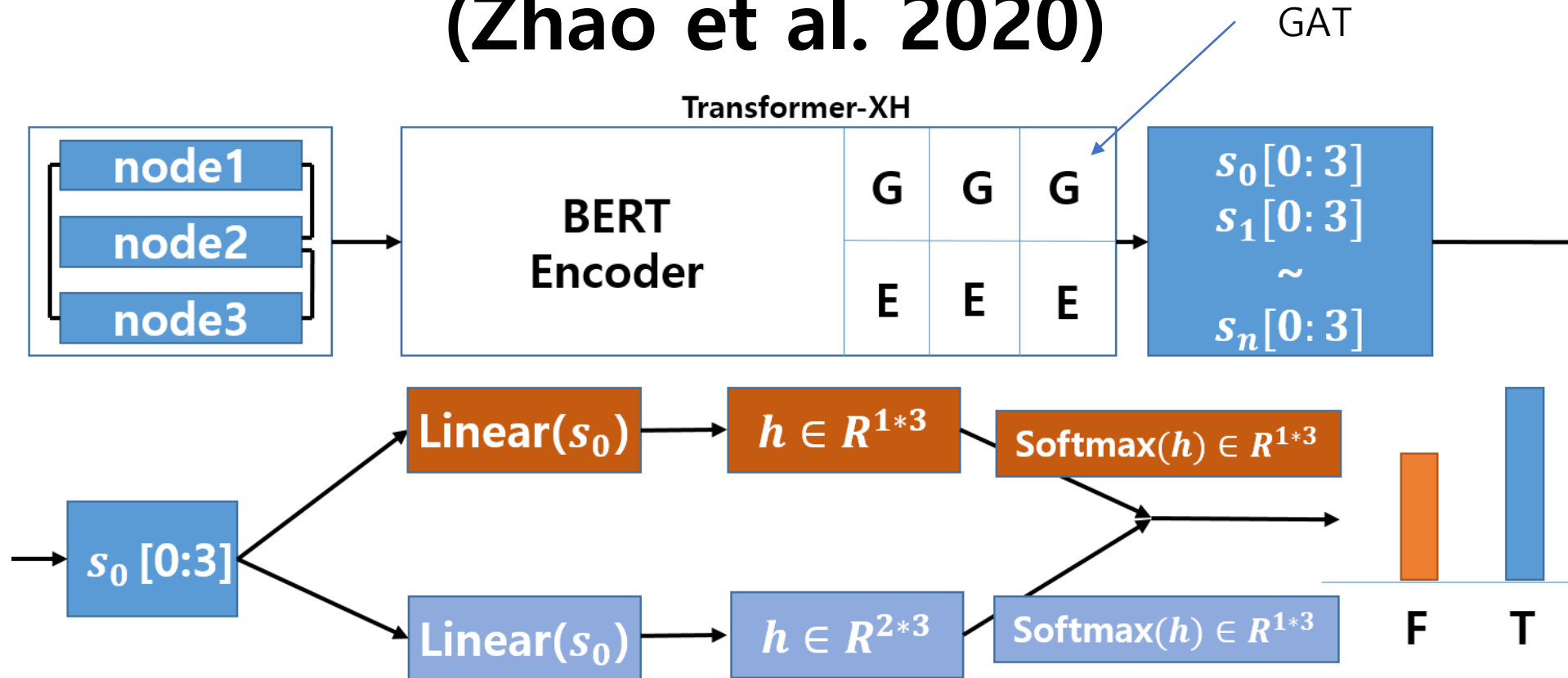
문장 유사도

증거 문장 선택



인공 신경망 증거와 질문 답변을 통한 사실 검증

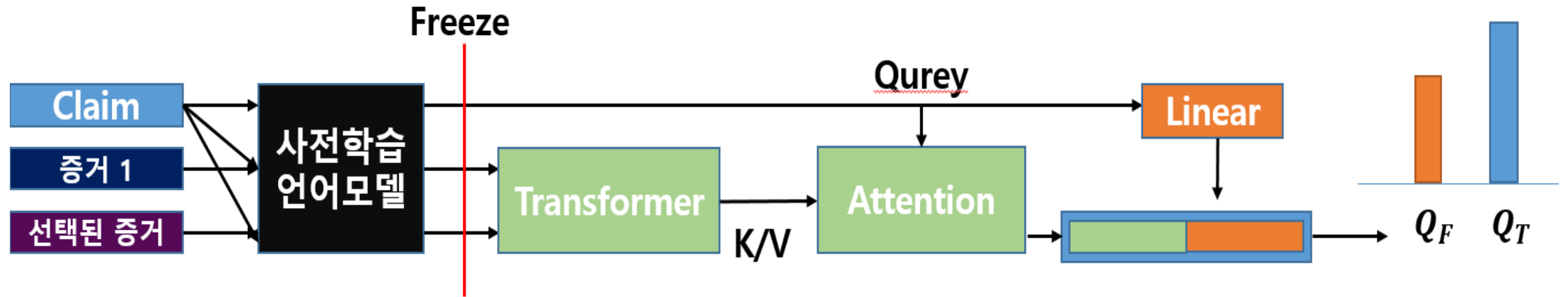
Transformer – XH (Zhao et al. 2020)



Node : $s+Q+s'+A+Title+s'+Sentence+s'$

s : 분류토큰 s' :문장 분리토큰

DQN 사실검증 (Wan et al. 2021)



증거 비교

질문 : 행복한 바오밥은 어디에 위치하고 있어?

답변 : 행복한 바오밥은 서울특별시 동작구 여의대방로 2길 22에 위치하고 있습니다.

TF-IDF

제목 : 행복한 바오밥

증거: 본사는 서울 특별시 마포구 토정로 47에 위치하고있다.

제목 : 의료법인 일심 재단

증거: 포천시 소흘읍 661번길에 위치하고 있다.

제목 : 행복한 바오밥

증거: 각각 도블,셈셈피자,...,을 판매하고 있다.

제목 : 행복한 바오밥

증거: 행복한 바오밥은 2004년 10월에
설립된 대한민국의 보드게임 회사이다

제목 : 사랑의 장기기증 운동 본부

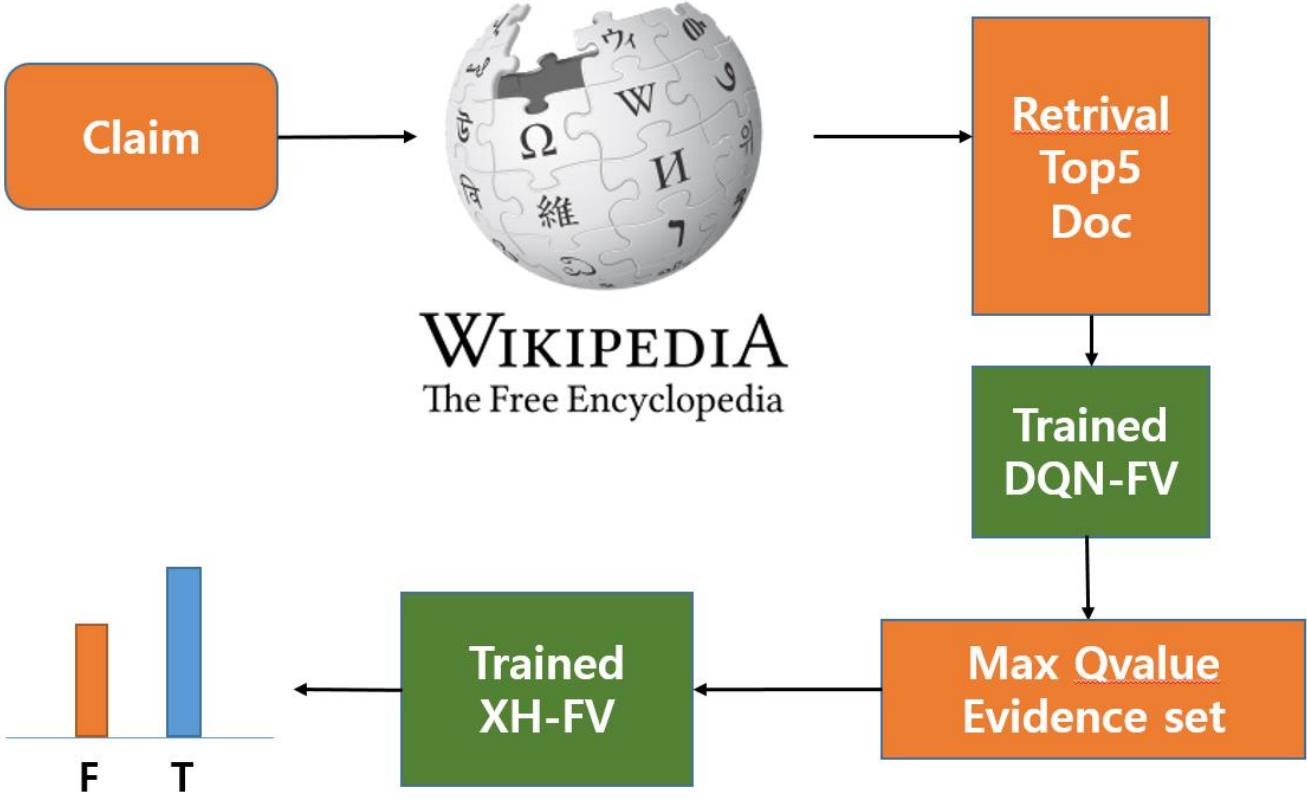
증거: 먼 훗날 세상을 떠나게 될 때에 장기를 기증하겠다는
약속의 의미로 만 16세 이상이라면 누구나 참여 가능합니다

DQN

제목 : 행복한 바오밥

증거: 본사는 서울 특별시 마포구 토정로 47에 위치하고있다.

두 모델의 결합



HyperCLOVA를 이용한 데이터 구축

증거에서 토픽 추출

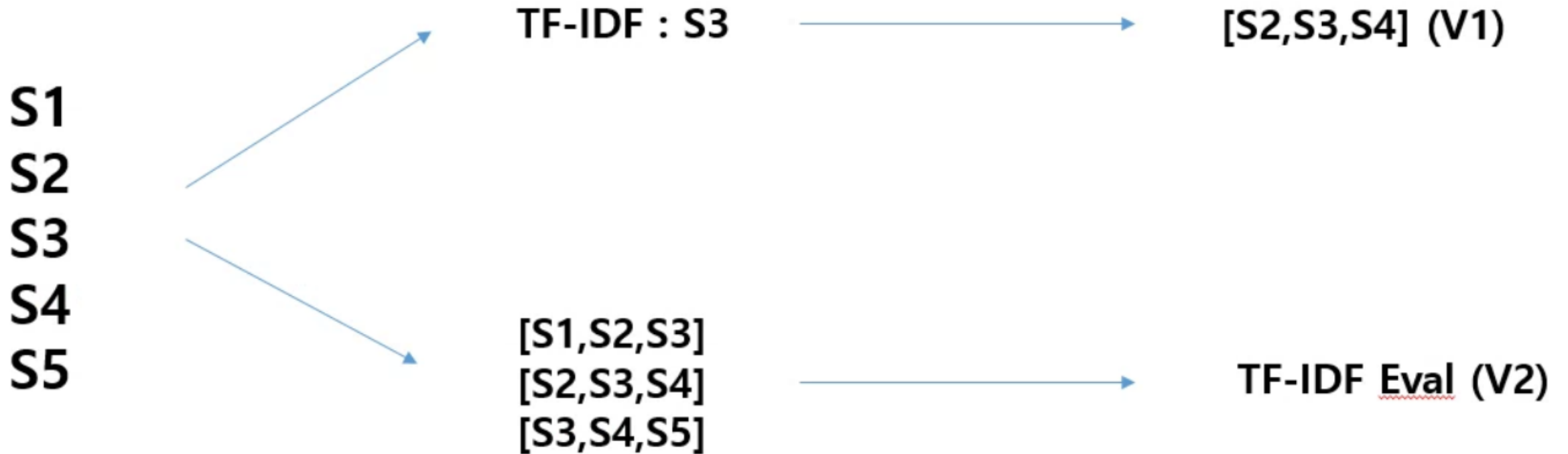
토픽을 통한 질문생성

질문과 증거에서 답변 생성

오답을 위해 답변 변형

입력	전라북도는 대한민국의 서남부에 있는 도이다.
출력	전라북도는 “대한민국의 서남부”에 있는 도이다.
입력	전라북도는 “대한민국의 서남부”에 있는 도이다.
출력	전라북도의 위치는 어디야?
입력	지문 : 전라북도는 대한민국의 서남부에 있는 도이다. 질문 : 전라북도의 위치는 어디야?
출력	전라북도의 위치는 대한민국 서남부 입니다.
입력	전라북도의 위치는 대한민국 서남부 입니다.
출력	전라북도의 위치는 대한민국 동북부 입니다.

추가 실험 : 문장의 단락 구성방법



실험 결과

표 2. 증거문 정밀도에 따른 정확도

XH	Dev
only Golden	93.80
2개	93.47
5개	93.66

**단순한 증거의 정밀도 향상은
성능향상을 보장하지 못함**

실험 결과

표 1. 정확도 측정 결과

	Dev	Test
XH	93.66	60.56
DQN	85.44	57.27
XH-DQN	90.97	60.09

실험 결과

표 2. 증거문 정밀도에 따른 정확도

XH	Dev
only Golden	93.80
2개	93.47
5개	93.66

**단순한 증거의 정밀도 향상은
성능향상을 보장하지 못함**

실험 결과

표 3. 문장 구성에 따른 정확도

	Dev	test
XH(문장)	92.89	60.09
XH(Ver1)	93.66	59.62
XH(Ver2)	92.13	62.44

**단락 구성 보다는
증거의 집중이 좋은 결과를
보여준다고 생각됨**