

# BERT 문장 임베딩과 Co-Attention 메커니즘 결합에 기반한 알츠하이머병 치매와 조현병 진단

정민교<sup>o</sup>, 나승훈,  
김고운, 신병수, 정영철  
Jungmk@jbnu.ac.kr<sup>o</sup>



Jeonbuk National University\*  
Jeonbuk National University Hospital

# 목 차

---

- 서론

- 치매 및 조현병
- 관련 연구

- BERT 문장 임베딩과 Co-Attention 메커니즘

- 결합 분류 모델

- 분류 과정

- 실험

- 데이터셋
- 실험 결과

- 결론 및 향후 연구

- 결론
- 향후 연구

# 서론

---

- 치매

: 기억, 언어, 판단력 등 여러 영역의 인지기능이 감소해 일상 생활에 지장을 초래할 수 있는 질병

- 세부적 분류

- ① 알츠하이머성 치매
- ② 경도인지장애 (Mild Cognitive Impairment)
- ③ 아미노이드 양성반응 치매 환자군

< 알츠하이머성 치매 환자군 >

의사: 어머니 지금이 몇 년도 인가요?

환자: 금방 생각이 안나요.

의사: 몇 년도?

환자: 1989년인가?

# 서론

---

- 조현병

: 와해된 언어의 사용과 잦은 망상으로 사회적 관계에 피해를 초래할 수 있는 중증 정신질환

- 세부적 분류

- ① 조현병
- ② 조현병 스펙트럼 (조현병, 분열정동장애, 정신 분열형 장애 등)
- ③ 인지기능 영역 결함 조현병

〈 조현병 환자군 〉

의사: 혼자 있는데 옆에서 모르는 사람들이 수군 거리고 떠들면 어떤 기분이 드세요?

환자: 괜히 불쾌합니다.

의사: 무엇 때문인지 말씀해 주세요.

환자: 절 얘기하는 것 같아 가지고 불쾌합니다.

# 서론

---

- 관련 연구

- ① 영어 데이터셋 대상 연구

- To bert or not to bert: Comparing speech and language-based approaches for alzheimer's disease detection [Balagopalan et al '20]

- : 치매 분류 성능에 대해 SVM 모델과 BERT 모델을 비교하여 BERT 모델의 F1 성능이 우수한 것을 확인

- 딥러닝 기법의 적용이 단순 기계학습 기법과 비교해 우수한 성능 달성

- ② 한국어 데이터셋 대상 연구

- BERT와 Co-Attention 메커니즘을 이용한 알츠하이머병 치매와 조현병 진단 [정민교 et al '21]

- : Co-Attention 메커니즘을 적용하여 치매와 조현병 분류 성능의 향상을 확인

- 데이터셋을 확장하여 질병의 세부적 분류를 수행하였고, Gating Mechanism을 사용

# BERT 문장 임베딩과 Co-Attention 메커니즘 결합에 기반한 알츠하이머병 치매와 조현병 진단

- 분류 과정

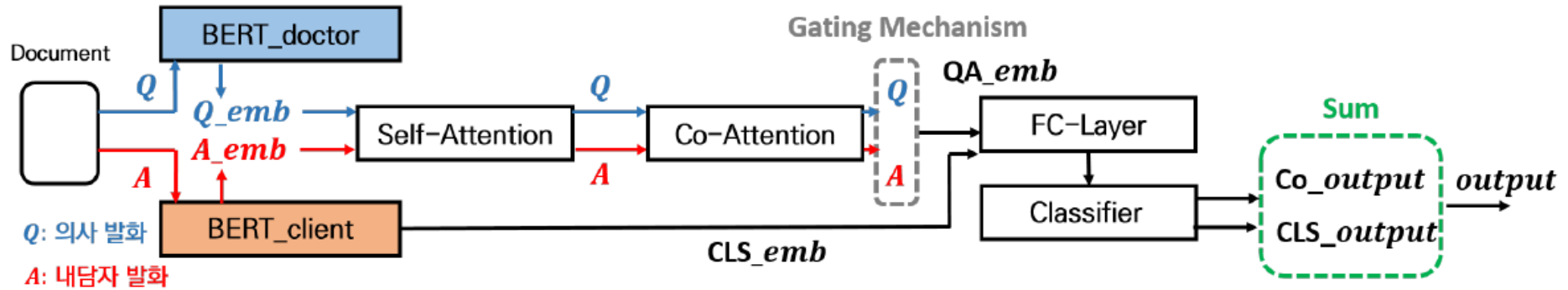


그림 1: Co-Attention + CLS 모델

① Co-Attention Mechanism

$$F = \tanh(Q^T W_l A) \quad (1) \quad a^a = \text{softmax}(w_{ha}^T H^a) \quad (4)$$

$$H^a = \tanh(W_a A + (W_q Q) F) \quad (2) \quad a^q = \text{softmax}(w_{hq}^T H^q) \quad (5)$$

$$H^q = \tanh(W_q Q + (W_a A) F^T) \quad (3) \quad \hat{a} = a^a A, \quad \hat{q} = a^q Q \quad (6)$$

$$QA_{emb} = \text{Gating}(\hat{a}, \hat{q}) \quad (7)$$

# BERT 문장 임베딩과 Co-Attention 메커니즘 결합에 기반한 알츠하이머병 치매와 조현병 진단

- 분류 과정

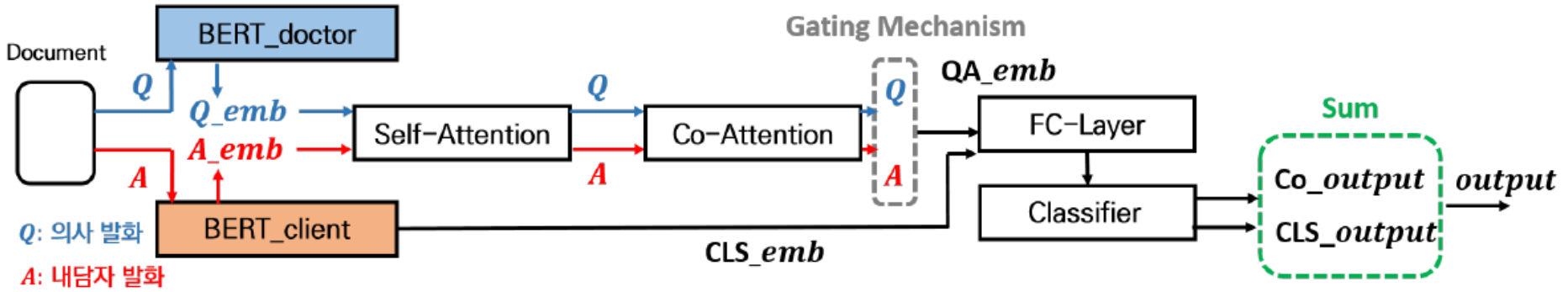


그림 1: Co-Attention + CLS 모델

## ② Gating Mechanism

$$gating_{ratio} = Sigmoid(Linear(\hat{a}) + Linear(\hat{q})) \quad (1)$$

$$QA_{emb} = (gating_{ratio} * \hat{a}) + ((1 - gating_{ratio}) * \hat{q}) \quad (2)$$

## ③ Co-Attention output + CLS token 결합

$$CO_{output} = Classifier(FC\_Layer(QA_{emb})) \quad (1)$$

$$CLS_{output} = Classifier(FC\_Layer(CLS_{emb})) \quad (2)$$

$$output = CO_{output} + CLS_{output} \quad (3)$$

# 실험

- 데이터셋

: 특정 프로토콜에 따라 의사의 질문과 내담자의 답변을 녹음한 음성 시료를 전사 작업한 스크립트 데이터셋을 실험에 사용, 학습 데이터(8) : 검증 데이터(1) : 평가데이터(1) 비율로 나누어 실험 진행

Task	Train data	Dev data	Test Data
알츠하이머성 치매 환자군 vs 정상군	38 : 161	5 : 20	5 : 20
경도인지장애 환자군 vs 정상군	116 : 161	15 : 20	15 : 20
알츠하이머성 치매 환자군 vs 경도인지장애 환자군 vs 정상군	38 : 116 : 161	5 : 15 : 20	5 : 15 : 20
아미노이드 양성 환자군 vs 아미노이드 음성 환자군	30 : 38	4 : 6	4 : 6

< 치매 관련 Task 데이터셋 >

Task	Train data	Dev data	Test data
조현병 환자군 vs 정상군	159 : 125	20 : 16	20 : 16
조현병 스펙트럼 환자군 vs 정상군	190 : 125	24 : 16	24 : 16
인지기능 결함(2개 이상) vs 결함 없는 환자군	85 : 71	11 : 9	11 : 9
인지기능 결함(전역적) vs 결함 없는 환자군	60 : 92	7 : 12	8 : 12

< 조현병 관련 Task 데이터셋 >



# 실험

---

- 실험 결과

- Baseline

- : BERT 임베딩에 대해 Co-Attention 메커니즘을 사용한 모델을 사용<sup>1)</sup>

- Setting

- ① seed 값을 바꾸어 3번의 실험 성능의 평균 값을 사용
    - ② 환자군 예측에 대한 F1-score를 사용

---

1) 정민교, 나승훈, 김고운, 신병수, and 정영철, "BERT와 Co-Attention 메커니즘을 이용한 알츠하이머병 치매와 조현병 진단". 한국정보과학회 학술발표 논문집, pp 314-316, 2021.

# 실험

- 실험 결과

- 치매 관련 Task 실험 결과

Task	베이스라인	제안 모델
알츠하이머성 치매 VS 정상군	93.27	86.67(-6.6)
경도인지장애 VS 정상군	53.66	<b>56.69(+3.03)</b>
아미노이드 양성 VS 아미노이드 음성	52.91	<b>57.14(+4.23)</b>
알츠하이머성 치매 VS 경도인지장애 VS 정상군	56.99	<b>57.89(+0.90)</b>

- ① 알츠하이머성 치매 vs 정상군 Task를 제외하고, 세 가지 Task에서 성능이 향상된 것을 확인
- ② 경도인지장애 vs 정상군 Task, 아미노이드 양성 vs 아미노이드 음성 Task에 나타난 환자군은 상대적으로 알츠하이머성 치매 환자군 보다 정상군에 가까운 인지능력을 가져 분류 성능이 낮은 것으로 판단
- ③ 알츠하이머성 치매 환자군의 적은 데이터 수로 알츠하이머성 치매 vs 경도인지장애 vs 정상군 Task의 성능이 낮은 것으로 판단

# 실험

- 실험 결과

- 조현병 관련 Task 실험 결과

Task	베이스라인	제안 모델
조현병 VS 정상군	77.95	75.58(-2.37)
조현병 스펙트럼 VS 정상군	79.98	78.98(-1.00)
인지기능 2개 영역 결합 VS 결합 없는 환자군	47.92	<b>68.08(+20.16)</b>
인지기능 전역적 결합 VS 결합 없는 환자군	53.09	<b>59.31(+6.22)</b>

- ① 인지기능 결합 관련 Task에서 성능이 향상된 것을 확인
- ② 조현병 vs 정상군 Task와 조현병 스펙트럼 vs 정상군 Task의 성능이 하락하였지만, 성능이 향상된 두 Task의 상승폭이 더 크기 때문에 유의미한 결과라 판단
- ③ 조현병 vs 정상군 Task와 조현병 스펙트럼 vs 정상군 Task의 확보한 데이터셋 수가 더 많아 베이스라인과 제안 모델 성능이 두 Task와 비교해 성능이 우수한 것으로 판단

# 실험

---

- 결론 및 향후 연구

- 결론

- ① gating mechanism을 통해 일부 Task에서 기존 Baseline과 비교해 성능 향상을 보임
    - ② 문장 분류에 사용되는 CLS 토큰의 사용으로 일부 Task의 성능 향상을 확인
    - ③ 치매 및 조현병 환자군의 세부적인 분류를 통해 다양한 분류 실험 수행

- 향후 연구

- ① 다양한 언어적 자질 활용 한계
      - ➔ 발성 중 환자의 pause 태그, 말 더듬 태그, 반복 발화 태그 등 스크립트에 표현된 언어 표현 태그 활용
    - ② 모델 예측에 대한 신뢰성 확보 한계
      - ➔ 어텐션 가중치 시각화 및 설명 가능한 인공지능 기법 적용을 통해 모델 신뢰성 확보

**Thank You**