**AI 기반 선박설계 관련 Q&A 보조 시스템 개발**

**(결과 보고서)**

           연구 기간 :  2021. 10. 01. ～ 2022. 09. 30.

           연구 책임자 : 한국과학기술원 수리과학과 교수 이창옥

           연구원: 한국과학기술원 수리과학과 교수 전현호, 대학원생 이영규, 김광우

**목 차**

**1. 연구과제명 ……………………………………………………………………………………………… 3**

**2. 연구기간 ………………………………………………………………………………………………… 3**

**3. 연구 배경 및 목표 …………………………………………………………………………………… 3**

**4. 연구개발 내용 ……………………………………………………………………………………………… 4**

**5. 참고문헌** ………………………………………**………**…**…………………………………**……………………… **5**

**1. 연구과제명 :** AI 기반 선박설계 관련 Q&A 보조 시스템 개발

**2. 연구기간 :** 2021. 10. 01. ～ 2022. 09. 30.

**3. 연구 배경 및 목표**

자연어 처리 (Natural Language Processing) 방법론의 발전으로, 문맥을 통한 단어의 이해가 가능하고, 이는 language translation, question and answering, sentiment analysis, chatbot, grammar check algorithm 등으로 광범위하게 이용되고 있다. 이러한 성공적인 사례에도 불구하고, 산업 현장의 수요를 맞추기에는 부족한 점들이 있어 여러 부분에서의 보완이 필요하다. 본 과제에서는 Attention Mechanism이라는 최신 방법을 적용하여 조선업 (shipbuilding industry)에 전문화된 언어 모델을 개발한다.

* 1. **Natural Language Processing and recurrent neural networks**

DSEC에서는 조선(shipbuilding)에 관한 질의/응답을 Plan Closed Form (PCF) 파일의 형태로 오랜 기간 저장해왔는데, 이를 자연어 처리(Natural Language Processing)를 통해 모형을 세워 활용하려고 한다. 단어의 영향력을 해석하는 전통적인 방법은 순환 신경망(Recurrent Neural Networks)이다. 하지만 순환 신경망은 긴 문장에 대한 정보 압축을 효과적으로 하지 못한다. 예를 들어, Long-Term Dependency 문제 때문에 긴 문장에 대해서 원하는 결과를 잘 반영하지 못할 수 있다. 이러한 문제를 능률적으로 해결한 것이 Attention Mechanism이다. 이를 통해 순환 신경망의 계산에 있어서의 비효율성 문제를 해결함과 동시에 성능도 월등히 개선할 수 있다.

* 1. **Attention Mechanism**

Attention Mechanism이란 각각의 데이터 처리 과정에서 기존의 문맥을 참고하되 가장 중요한 의미와 위치에 주목하는 알고리즘이다. 이는 가변적인 문장에 대해 유동적으로 압축을 하여 정보 손실이 적고, 기울기 소실 문제를 해결했으며, 각 단어의 위치를 문맥에 따라 가중치를 잘 반영한다. 하지만 Attention Mechanism을 이용한 Transformer 구조의 근본적인 목적은 언어 번역이다. 우리의 목적에 알맞게변형하려면 이 구조의 input embedding에 집중할 필요가 있다. 이러한 관점을 잘 반영한 모델이 BERT이다. 이 모델은 bidirectional method, pre-training, fine-tuning, masked language model과 같은 아이디어를 접목하여 embedding 성능을 더욱 향상시켰다.

**4. 연구개발 내용**

**4.1 Fine-tuning for Bidirectional Encoder Representations from Transformers (BERT)**

BERT는 인간의 언어를 이해하는 Language model이다. 언어를 학습하기 위해선, 방대한 양의 학습자료를 필요로 하기 때문에 dsec의 PCF 자료만으로 BERT를 training 하는 것은 어렵다. 따라서 대량의 데이터에서 학습이 완료된 BERT로부터 dsec의 데이터를 학습하는 Transfer learning 방법으로 BERT를 개발하였다. 즉, 사전 훈련된 BERT를 이용하여 새로운 dsec의 데이터를 Masking 알고리즘을 통해 Fine-tuning 시켰다. 아래 그림은 BERT의 학습 과정에 대한 모식도이다. 학습 후, BERT는 DB를 적절한 Vector space에 Embedding을 할 수 있다. 따라서, 입력 값을 BERT를 통해 Embedding 한다면 DB와 유사도 비교를 통해 Information retrieval을 수행할 수 있다.

**4.2 Okapi BM25**

Information retrieval을 효과적으로 수행하기 위해선, document를 사전 indexing 하여 word matching system을 구축 해 놓는 것이 효과적이다. 왜냐하면 DB가 클수록 Inference 시간이 길어 질 수 있기 때문이다. Okapi BM25 (이하 BM25)는 이를 수행할 수 있는 전통적인 Machine learning model중 하나이다. 이 모델은 inference task의 속도를 향상시킬 뿐만 아니라 BERT의 정확도까지 향상시킬 수 있다. 구체적으로, BM25로 관련 있는 것으로 추정된 DB를 우선적으로 추려낸 뒤, BERT의 유사도 비교를 수행하였다. 더욱이, BM25를 통해 ranking 시스템을 개발하여 BERT와 융합한 것이 최종적으로 가장 좋은 성능을 냈다. 이 과정에서 FLAIR 라는 pretrained deep learning model이 word (or sentence) tokenizer와 part-of-speech tagger을 제공함으로써 융합과정에 도움을 주었다.

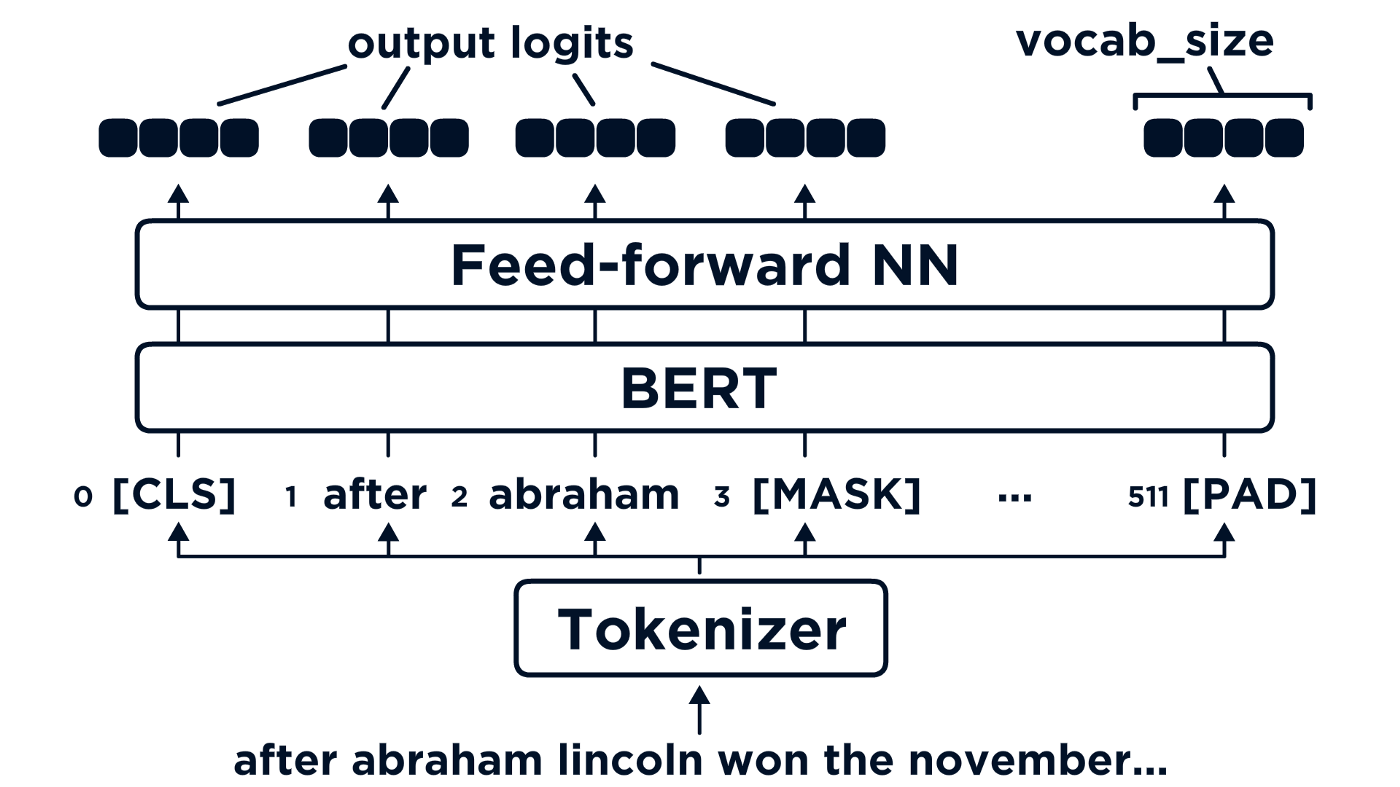


Figure 1. BERT Training의 모식도

**4.3 Software implementation**

개발된 결과물은 실사용자가 사용하기 쉬운 소프트웨어의 형태로 packaging 했다. 이를 통해 Windows 환경에서 실 사용과 학습 모두 가능케 했다. 또한, 사용자 매뉴얼을 제작하여 사용자의 효과적인 사용을 기대할 수 있게끔 했다. 매뉴얼은 첨부파일에서 확인할 수 있다.

검색 측면에서, Sentence search는 주로 연구한 information retrieval 방법이다. 사용자의 질의와 가장 유사한 형태를 가진 문장을 DB에서 찾으며 BERT, BM25, FLAIR와 같은 알고리즘이 적절하게 동시 작동한다. 좀 더 풍부한 활용을 위해서, Keyword search, 확장 검색, 결과물 저장과 같은 기능이 제공된다.

학습 측면에서, 개발자가 아니더라도 학습시킬 수 있게끔 GUI 환경을 개발했다. GUI는 새로운 데이터를 DB에 추가하여 검색 결과를 더욱 풍성하게 만들고 AI 모델을 학습시킴으로써 언어 이해 능력을 더욱 향상시킬 수 있다. 게다가, 학습 방법을 Full training과 Instant training으로 구별함으로써 사용자의 편의를 더욱 증진시켰다. 자세한 두 학습 방법의 장단점은 사용자 매뉴얼에서 찾을 수 있다.

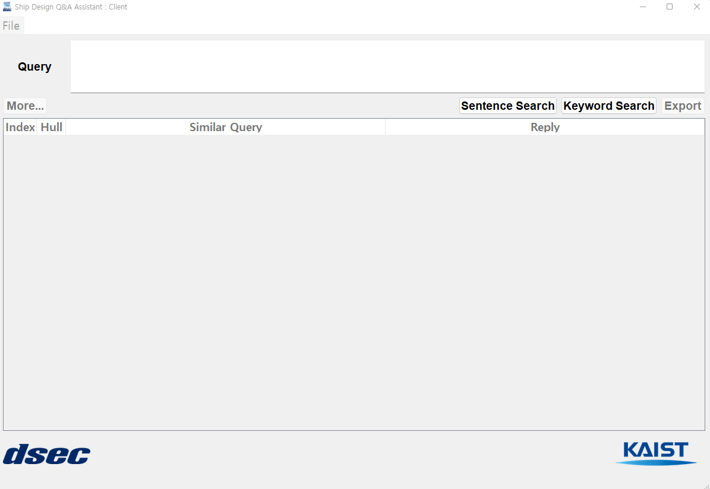


Figure 2. 실행화면

사용자 매뉴얼: 

**5. 참고문헌**

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 5998-6008).

[2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

[3] Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., & Vollgraf, R. (2019, June). FLAIR: An easy-to-use framework for state-of-the-art NLP. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics (demonstrations)* (pp. 54-59).

[4] Trotman, A., Puurula, A., & Burgess, B. (2014, November). Improvements to BM25 and language models examined. In *Proceedings of the 2014 Australasian Document Computing Symposium* (pp. 58-65).