

Baseline of SWOT Classification using Bidirectional Encoder Representations from Transformers for Business Intelligence Cloud Platform

Okestro

Yongshin Kim

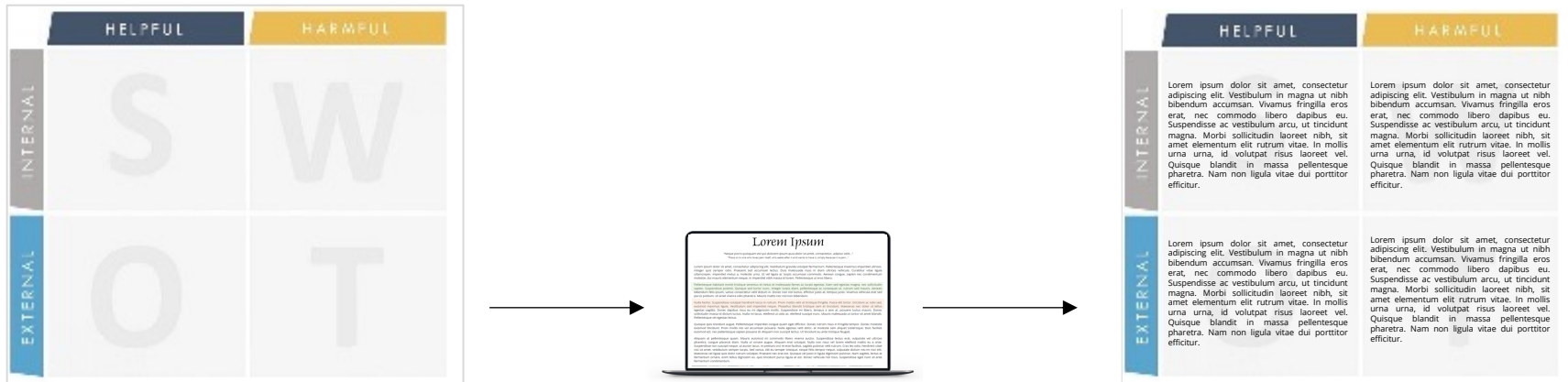


※ This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00147, A Technology of Integrated Management Platform for Multi-Hybrid SaaS Solution)



- The Business Intelligence system is a **decision support system** to provide key information about the company.
- Establishing the right business strategy through the analysis of **internal** and **external** environmental factors of a company can lead to corporate innovation and performance creation.
- Therefore, the process of recognizing the **strengths** and **weaknesses** of the company and accurately analyzing the **opportunities** and **threats** outside the company is critical.

Introduction



- As part of the development of marketing artificial intelligence services that can be commercialized in cloud-based platforms, this study proposes a method of automating **SWOT classification** using artificial intelligence.

Problem Setup

1) There is not much text data labeled with S/W/O/T.

www.jobkorea.co.kr

www.catch.co.kr

www.marketing91.com/

...



**Not enough number to
train deep learning model**



Problem Setup

1) There is not much text data labeled with S/W/O/T.

www.jobkorea.co.kr
www.catch.co.kr
www.marketing91.com/
...



**Not enough number to
train deep learning model**

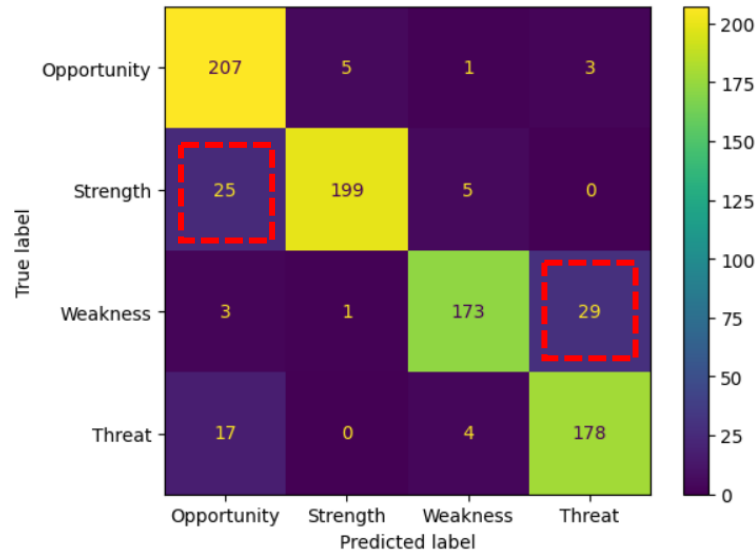


Sol) BERT

- SWOT classification through **Bidirectional Encoder Representations from Transformers(BERT)**, which is widely used in the natural language processing.
- Transfer learning
 - **Storing** knowledge gained while solving one problem and **applying it to a different but related problem.**

Problem Setup

2) The model cannot distinguish between
S and O, W and T.



Strength? Opportunity?

- “이 기세를 몰아 하이퍼넥트는 독자적인 기술을 바탕으로 북미 독일 지역에 신규 서비스를 출시하며 지속적으로 해외 시장을 확대 공략 중이다.”

Weakness? Threat?

- “엔씨소프트는 리니지와 같은 하드코어 장르에 특화된 게임사로서, 캐주얼 장르에서 큰 성공을 거둔 경험이 없다. 특히 리니지 하면 떠오르는 부정적인 이미지로는 ‘린저씨’, ‘과도한 과금 유도’ 등이 있다. 이러한 하드코어 RPG는 최근 새로운 소비 주축으로 떠오른 MZ세대에게 큰 인기를 끌지 못하고 있다.”

Problem Setup



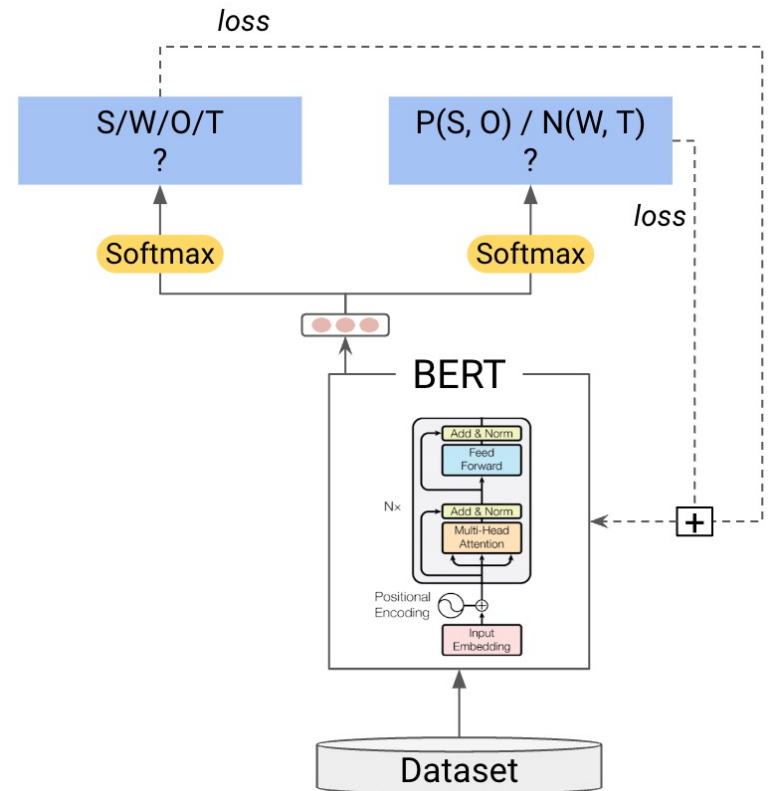
2) The model cannot distinguish between
S and O, W and T.

Sol) Multitask Learning

- Multiple learning tasks are solved **at the same time**, while exploiting commonalities and differences across tasks.

Overall Architecture

- BERT model that uses only the **encoder** portion of the transformer.
- Models learned with **less data** using **pretrained model** on a hugging face.
- Multitask model consisting of a task that classifies **S/W/O/T** and a task that classifies **positive(S, O)** and **negative(W, T)**.
 - By **combining the loss values** extracted from each of the two tasks, the model will be able to distinguish **S and O, W and T** better.



- The dataset consists of a text paragraph containing each S, W, O, and T content for the company and a corresponding label (S, W, O, and T) for each paragraph.
- The total number of paragraphs and corresponding labels is **5,650**, and the total number of sentences is **28,591**.
- Korean Language Understanding Evaluation(KLUE) pretrained language model.
- Hyper-parameter
 - Learning rate: 0.0005
 - Warmup ratio: 0.2
 - Epoch: 20
 - Batch size: 8, 32, 64

Experiment Results

Type	Batch size	Accuracy	Precision	Recall	f1 score
Paragraph	8	0.907	0.907	0.904	0.905
	32	0.901	0.902	0.901	0.901
	64	0.907	0.908	0.905	0.905
Sentence	8	0.741	0.739	0.738	0.737
	32	0.742	0.739	0.738	0.738
	64	0.747	0.743	0.74	0.74

- Even with the use of the basic BERT model, paragraph version were able to have an accuracy of about **90%**.
- The accuracy of the sentences seemed to be **limited to some extent** because they had **sequential dependency**.

- This study classified **SWOT**, one of the commonly used analysis tools in the environmental analysis process of companies, using **BERT**, which is widely used in the field of natural language processing.
- Contribution of this study is that the **baseline** model is established for the first time by incorporating AI that has not been attempted so far into SWOT analysis.
- Discussion
 - How can we consider the **sequential dependency** between sentences when using sentence version?
 - If the **subject in the sentence changes**, the answer may change directly from "threat" to "opportunity." How can a model learn these things well?

AI in BI

Strength

"오케스트라는 클라우드 데이터센터를 구축하기 위한 **A부터 Z까지 풀스택 솔루션을 모두 보유하고 있다**. 김 총괄대표는 **"글로벌 외산 제품도 구현하지 못한 기능들을 독자 기술력으로 구현해 국내 뿐 아니라 많은 국제 특허도 보유하고 있다"**며 기술력에 자신감을 보였다."

91%

Weakness

앞서 이 부회장은 2019년 "메모리에 이어 파운드리를 포함한 시스템 반도체 분야에서도 확실히 1등을 하겠다"는 비전을 제시한 바 있다. 하지만 아직 제대로 된 성과는 내지 못했고, 오히려 격차가 벌어지는 상황이다. **삼성전자가 자랑해온 스마트폰 애플리케이션 프로세서(AP) '엑시노스 2200'은 올 초 낮은 성능과 수율 등의 문제가 발생**했고, 지난 4월 반도체 부문의 한 연구원이 이 부회장 등에게 전자우편을 보내 **경영진의 책임 회피와 패배의식 등 조직문화의 문제점을** 제기하기도 했다.

83%

AI SWOT Analysis Sample Based on Press Releases

Opportunity

"보고서를 작성한 강준영 교수는 **"반도체와 같이 대규모 투자와 연구개발에 오랜 시간이 소요되는 분야의 경우 정부가 인력·R&D·세제 등 전 분야에 걸쳐 연계하고 세밀하게 지원하는 게 필수적"**이라고 말했다. 이어 **"핵심 기술인력 확보의 경우 국내 우수인력 육성과 해외 핵심인력 유치를 동시에 진행하고 있다"**며 **"한국이 정책 활용 차원에서 검토해볼 수 있을 것"**이라고 했다.

86%

Threat

지난해까지 **코로나19 팬데믹**으로 급증했던 **반도체 시장 수요가 올 들어 둔화 조짐**을 보이는 데다 **전 세계적인 고물가·고금리·고환율의 3고(高) 악재**가 겹치는 상황과 맞물려 **모바일과 PC 등 전방 수요가 둔화되면서 메모리 반도체 시장이 조정 국면에 들어갈 것**이라는 분석이다. 특히 서버수요도 **글로벌 경기 위축 여파**로 줄어들 가능성까지 제기된 상태다.

91%

- [1] Hwang, J., Seo, J. H., Lim, J. S., Yoo, H. S., Park, J., Kim, Y. E., & Kim, J. H. (2017). Design and Implementation of Web-based SWOT Analysis Supporting Tool. *The Journal of the Korea Contents Association*, 17(7), 1-11.
- [2] Ying, Y. (2010, May). SWOT-TOPSIS integration method for strategic decision. In *2010 International Conference on E-Business and E-Government* (pp. 1575-1578). IEEE.
- [3] GURL, E. (2017). SWOT analysis: A theoretical review.
- [4] 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*.
- [5] Park, S., Moon, J., Kim, S., Cho, W. I., Han, J., Park, J., ... & Cho, K. (2021). Klue: Korean language understanding evaluation. *arXiv preprint arXiv:2105.09680*.
- [6] Caruana, R. (1997). Multitask learning. *Machine learning*, 28(1), 41-75.