

**BERTopic과 LLM(Large Language Model)을 결합한
경영학 학술연구 동향 분석 사례연구**

<목차>

요 약	1
I. 서론	2
1. 사례 연구 개요	2
2. 분석 개요	2
3. 연구의 기대효과	2
II. 사례 연구 설명	3
III. 분석 방법론	4
1. 과학계량학(Scientometrics)	4
2. 과학계량학에서 토픽 모델링의 활용 이유와 BERTopic	5
3. 토픽 모델링 결과에 대한 대규모언어모델 기반 라벨링 적용	7
IV. 데이터 분석	9
1. 분석 개요	10
2. 사례 분석: 경영학 분야의 학술연구 동향 분석	11
V. 결론	27
1. 시사점	27
2. 한계점	27
참고자료	28
Appendix	29

<표 목차>

<표 1> 소거 저널 목록.....	12
<표 2> 대규모언어모델 기반 토픽 라벨링 결과 일부.....	15
<표 3> 토픽 2 명명을 위해 대규모언어모델에 주입된 논문 10개 예시	16

<그림 목차>

<그림 1> Web of Science에서 검색한 LDA 논문 및 BERTopic 논문의 트렌드	6
<그림 2> BERTopic의 모듈성 (참고: Grootendorst).....	7
<그림 3> 분석 절차.....	9
<그림 4> BERTopic visualize_hieararchy 수행 결과.....	14
<그림 5> 경영학 학술 출판물의 양적 성장과 인용 동향: 2000-2023년 분석	17
<그림 6> 경영학 분야 주요 저널의 출판물 수 동향	18
<그림 7> 경영학 분야 주요 저널의 출판물 수 동향 Citation Topic Macro Level 분석	19
<그림 8> 경영학 연구의 주제별 출판 동향: Citation Topic Meso Level 분석	20
<그림 9> 경영학 연구의 주제별 출판 동향: Citation Topic Micro Level 분석	22
<그림 10> 경영학 연구 토픽(408개)의 진화: 2000-2023년.....	23
<그림 11> 출판 성장률을 고려한 주요 토픽 막대 그래프.....	24
<그림 12> 주요 핫 토픽에 대한 산점도.....	25

<요약>

경영학은 조직의 운영, 전략, 혁신 등 다양한 주제를 다루는 학문으로, 사회와 경제의 변화에 따라 지속적으로 진화하고 있다. 이러한 경영학 분야의 동향을 파악하는 것은 연구자, 실무자, 정책 입안자 모두에게 중요한 과제이다. 그러나 방대한 양의 학술 문헌과 데이터 속에서 유의미한 정보를 추출하고, 이를 바탕으로 연구 트렌드를 분석하는 일은 매우 어렵고 시간이 많이 소요된다.

이에 따라, 최신 자연어 처리(NLP) 기술을 활용한 자동화된 분석 방법의 필요성이 증가하고 있다. 특히, BERTopic과 대형 언어 모델(LLM: Large Language Model)의 결합은 방대한 텍스트 데이터를 효율적으로 처리하고, 심층적인 주제 분석을 가능하게 함으로써, 보다 객관적이고 정교한 연구 동향 분석을 수행할 수 있다.

특히 BERTopic은 비지도 학습 기반의 토픽 모델링 기법으로, 문서 집합에서 주제를 자동으로 식별하고 구조화된 방식으로 시각화할 수 있는 강력한 도구이다. 문서의 문맥을 이해하는 데 뛰어난 성능을 보이는 임베딩 기법을 적용할 수 있어, 보다 효과적인 토픽 모델링을 가능하게 만든 도구이다. 또한, 토픽 모델링(Topic Modeling)은 문서 집합에서 주제를 자동으로 식별하고 분류하는 중요한 기법이다. 그러나 토픽 모델링의 결과를 해석하고 의미 있는 라벨을 붙이는 과정은 여전히 도전적이다. 이러한 라벨링 과정에 대규모 언어모델(LLM: Large Language Model)을 사용하는 것은 정확도 향상과 효율성을 극대화할 수 있다.

본 연구는 이러한 필요성에 부응하여, BERTopic과 LLM을 결합한 방법론을 제시하고, 이를 통해 경영학 분야의 최신 연구 동향을 체계적으로 분석하고자 한다. 이를 통해 경영학 연구자들은 보다 정확하고 신뢰할 수 있는 정보를 바탕으로 연구를 수행할 수 있으며, 나아가, 본 연구는 다른 학문 분야에서도 유사한 방법론을 적용하여 연구 동향을 분석하는 데 기여할 수 있을 것으로 기대된다.