

Capital Market 에서 LLM 기반의 포트폴리오 자산운용 자동 결과 생성

이건무¹, 권윤재¹, 오정은¹, 류다현¹, 장준혁¹
¹서경대학교 컴퓨터공학과

genmu4@skuniv.ac.kr, yjyjk0127@gmail.com, jungeuno76@skuniv.ac.kr,
fbekgus413@naver.com, joonhyeok426624@gmail.com

Generating automatic results of LLM-based portfolio asset management in Capital Market

Gun-Mu Lee¹, Yun-jae Kwon¹, Jung-Eun Oh¹, Da-hyun Ryu¹, Jun-Hyuk Jang¹
¹Dept. of Computer Engineering, SEOKYEONG University

요 약

본 연구는 Capital Market 에서 LLM 기반의 포트폴리오 자산운용 중 ETF, MMF, TD, 자산배분형 상품, 채권형펀드 등 고객에게 제공하는 자산운용 보고서를 자동으로 생성하는 시스템을 개발하고 평가했다. GPT-4, M2M-100, KoBART 모델을 결합해 번역, 요약, 보고서 작성 과정을 자동화했으며, 생성된 보고서와 원문 자산운용 보고서 간 평균 코사인 유사도는 90% 이상을 기록했다. 이 AI 시스템은 기존 방식 대비 효율성을 향상시켰으며, 금융 분야에서의 실용적 가치를 확인하였다.

1. 서론

글로벌 자산운용시장은 AI 발전과 디지털화의 가속화로 빠르게 변화하고 있으며, 이러한 변화는 2024년 우리나라 자산운용 산업에도 큰 영향을 미칠 것으로 예상된다. 펀드 판매의 온라인화, ETF의 진화, 그리고 자산관리 서비스의 대중화는 시장의 구조적 변화를 주도하고 있다.

자산운용보고서는 펀드의 운용 현황과 성과를 평가하고, 투자자에게 운용 전략과 경제 상황에 대한 정보를 제공하기 위해 작성된다. 보통 월간, 분기별 또는 연간 주기로 작성되며, 경제적 변수와 투자 성과의 상관관계를 분석해 향후 전략을 제시한다. 전통적으로 이 작업은 5년 이상의 경력을 지닌 펀드매니저가 수행하며, 보고서 작성에는 약 15일이 소요된다.

이러한 배경에서 LLM 기반의 자산운용보고서 자동 생성 기술을 제안한다. 이 기술을 통해 초급 펀드매니저도 효율적으로 보고서를 작성할 수 있으며, 숙련된 매니저에게는 업무 부담을 줄여 더 정교한 투자 전략 수립에 집중할 수 있는 환경을 제공한다.

2. 연구 방법

2.1 데이터 수집 및 전처리

자산운용 보고서 자동 생성 시스템 구축을 위해 경제 뉴스, 펀드 정보, 자산운용 보고서 등의 금융 데이

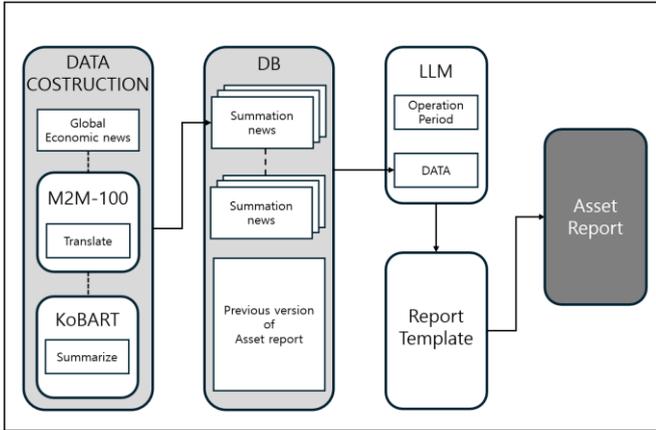
터를 수집하고 전처리했다. 국내 경제 뉴스는 네이버 등 포털 사이트에서, 해외 경제 뉴스는 Investing.com 등에서 크롤링하여 수집했다. 해외 뉴스는 Google API를 통해 번역되었으며, M2M-100 모델[1]의 번역 성능을 향상시키기 위한 자료로 사용되었다. 수집된 뉴스는 불필요한 정보를 제거한 후 금융 데이터에 최적화되었다. 번역된 기사와 원문을 KoBART 모델[2]로 요약하여 데이터베이스에 저장하고, 실시간으로 자산운용 보고서 작성에 활용하였다.

2.2 생성형 AI 모델의 학습

GPT-4o를 기반으로 M2M-100과 KoBART 모델을 결합하여 자산운용 보고서 자동 생성 시스템을 구축하였다. M2M-100은 다국어 번역 모델로, 해외 경제 기사를 한국어로 번역하는 데 사용되었다. 번역된 기사는 보고서 작성의 주요 자료로 활용되었으며, 번역 성능은 BLEU 스코어로 평가되었다. KoBART는 금융 기사 요약을 담당하는 한국어 텍스트 요약 모델이다. 약 25,000개의 데이터셋을 이용해 Fine-tuning을 거쳐 경제 데이터에 최적화되었으며, 이를 통해 자동 생성된 보고서의 품질이 크게 향상되었다. GPT-4o 모델은 M2M-100과 KoBART 모델을 통해 수집된 번역 및 요약 데이터를 바탕으로 자산운용 보고서를 자동으로 작성한다. 보고서 작성 과정에서는 실시간 경제 지표

와 자산운용 데이터를 반영하며, 사용자가 입력한 펀드명, 운용 기간, 수익률 등의 정보를 바탕으로 운용 보고와 향후 운용 계획을 자동으로 작성한다.

2.3 시스템 구성도

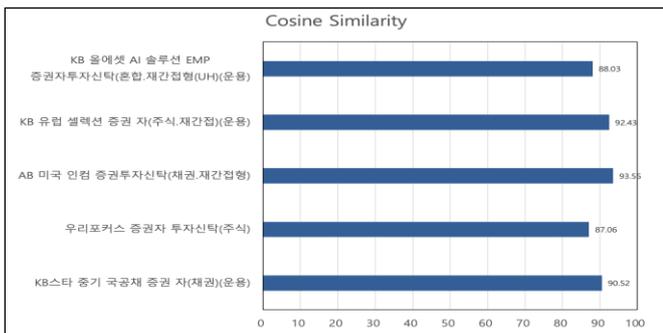


(그림 1) 시스템 구성도

(그림 1)은 자산운용 보고서 자동 생성 시스템의 구성도를 나타낸다. 데이터 구축 과정은 2.1 장에서 설명한 내용을 따르며, 생성된 데이터와 기존 자산운용 보고서 데이터는 LLM 모델(GPT-4o)의 프롬프트로 활용된다. 이를 기반으로 LLM 모델은 자산운용 보고서를 생성하며, 결과물은 작성된 보고서 템플릿에 적용되어 최종 자산운용 보고서가 자동으로 완성된다.

3. 실험 결과

3.1 자산운용 보고서 자동 생성 평가



(그림 2) 원문과의 운용보고 코사인유사도

(그림 2)는 AI 모델이 자산운용 보고서를 자동 생성하는 능력을 코사인 유사도로 평가하였다. 자산운용 사로부터 제공받은 10 개의 펀드 데이터를 바탕으로 AI 는 각 펀드의 기본 정보를 분석하고 운용 보고 및 향후 계획을 작성하였다. 생성된 보고서를 KoBERT 모델[3]을 통해 임베딩을 진행한 후 원문과 코사인 유사도[4]를 측정 한 실험 결과, AI 모델은 평균 90% 이상의 유사도를 기록하였다.

3.2 번역 및 요약 성능 평가

M2M-100 모델과 KoBART 모델의 번역 및 요약 성능을 평가하였다. M2M-100 모델은 74.42%의 BLEU 스코어로, Papago 와 Google Translate 보다 높은 번역 정확도를 기록하였으며, KoBART 모델은 일관성 있는 요약 결과를 제공하였다. 번역 및 요약된 데이터는 자산운용 보고서 작성에 실시간으로 효과적으로 반영되었다.

Model	ROUGE-1	ROUGE-2	ROUGE-L
T5	43.1	20.7	40.2
BART	44.1	21.3	40.9
Fine-Tuned KOBART	46.3	24.1	45.1

<표 1> 요약 모델 성능 비교(ROUGH Score)

<표 1>은 요약 모델들의 ROUGE 성능을 비교한 결과를 보여준다. 다른 모델들과 비교했을 때 KoBART 모델은 한국어 요약 작업에서 다른 모델들보다 높은 성능을 나타냈다.

4. 연구결과

본 연구에서는 LLM 을 활용해 자산운용보고서 자동화 서비스를 구현하였다. 웹 크롤링으로 수집한 최신 경제 뉴스를 M2M-100, KoBART 모델을 통해 번역 및 요약하여, 보고서 작성에 필요한 데이터의 신속한 반영이 가능해졌다.

이번 서비스의 도입으로 보고서 생성 시간 단축, 정확한 번역과 분석 제공, 그리고 경제 상황 반영이 가능해졌으며, 인력과 비용 절감, 고객 만족도 증대와 같은 실질적인 효과를 기대할 수 있다.

향후 연구에서는 BERT 와 같은 최신 언어 모델의 도입과 고품질 데이터셋 확보를 통해 모델 성능을 개선할 필요가 있다. 이를 통해 더 정교한 분석과 예측이 가능해지며, 자산운용보고서 자동화 서비스가 비즈니스 현장에서 더욱 큰 가치를 발휘할 수 있을 것으로 기대된다.

참고문헌

- [1] A. Fan, S. Bhosale, H. Schwenk, Z. Ma, A. El-Kishky, S. Goyal, M. Baines, O. Celebi, G. Wenzek, V. Chaudhary, N. Goyal, T. Birch, V. Liptchinsky, S. Edunov, E. Grave, M. Auli, and A. Joulin, "Beyond English-centric multilingual machine translation," *J. Mach. Learn. Res.*, vol. 22, pp. 1-48, 2021.
- [2] "KoBART", Github, last modified, Jun. 2022. <https://github.com/SKT-AI/KoBART> [accessed: Nov. 04, 2022]
- [3] "KoBERT", Github, last modified, Aug. 2022 <https://github.com/SKTBrain/KoBERT>
- [4] A. Huang, "Similarity measures for text document clustering," in *Proc. 6th New Zealand Comput. Sci. Res. Student Conf.*, Christchurch, New Zealand, 2008, pp. 49-56.