

# 다국어 AI 생성 뉴스와 사람 작성 뉴스의 언어적 특성 비교 및 판별 가능성 분석

최찬기<sup>†</sup>, 박채영<sup>‡</sup>, 김인우<sup>‡</sup>, 강채원, 한진영\*

성균관대학교

{<sup>†</sup>cck1216, <sup>‡</sup>codud0905, <sup>‡</sup>klwbty, codnjs3}@g.skku.edu, \*jinyounghan@skku.edu

## A Comparative Analysis of Linguistic Characteristics and Detectability between Multilingual AI-Generated and Human-Written News

Changi Choi, Chaeyoung Park, Linwoo Kim, Chaewon Kang, Jinyoung Han\*

Sungkyunkwan Univ.

### 요약

대규모 언어모델(LLM)의 발전으로 AI가 생성한 뉴스가 급증하며 정보의 출처와 신뢰도를 객관적으로 판단하는 기술의 필요성이 대두되고 있다. 본 연구는 AI가 생성한 뉴스와 사람이 작성한 뉴스 간의 언어적 특성 차이를 규명하고 이러한 차이가 탐지 성능에 미치는 영향을 분석하고자 한다. 영문 데이터를 다루는 기존 연구와 달리 국문 데이터를 포함한 다국어 AI 생성 뉴스 데이터셋을 구축하고, 이를 활용하여 문체·어휘·구조 측면에서의 정량적 분석을 수행한다. 분석 결과 AI 생성 및 사람 작성 뉴스는 정량지표 상 유의한 차이를 보였으며, 국문 뉴스는 영문에 비해 AI와 사람 작성 뉴스 간 변형 폭이 좁게 나타났다. 또한 LLM 기반 탐지 실험을 통해 분석 내용과 실제 탐지 성능이 일치함을 확인했다.

### I. 서 론

통신 기술의 발전에 따라 아날로그 신문이 아닌 인터넷 기사를 기반으로 한 신속한 정보의 확산 및 소비를 가능케 했다. 최근 대규모 언어 모델(Large Language Model, LLM)과 같은 딥러닝 기술이 뉴스 기사 생성에 활용되며 온라인 기사 생태계의 발전을 이룩함과 동시에 새로운 위기를 맞이하고 있다[1]. 사실 확인 절차가 누락된 AI 생성 뉴스의 급증에 따라 독자들은 정보의 출처와 신뢰도를 판단하기 어려워지면서, 이러한 뉴스의 특성에 대한 객관적 이해력 및 활용력이 강조되고 있다. 이를 위해 AI가 생성한 텍스트와 인간이 작성한 텍스트의 언어적 특성을 분석하여 그 차이를 파악하려는 연구들이 활발히 진행되어 왔으나, 대부분 영문 데이터에 국한되어 있다는 한계를 가진다[2]. 따라서 본 논문에서는 기존 연구의 범위를 확장하여 국문을 포함한 뉴스 데이터를 기반으로 AI와 사람이 작성한 텍스트의 언어적 특성을 심층적으로 분석하고 그 차이를 검증하고자 한다.

### II. AI 생성 뉴스 데이터셋 구축

영문 뉴스는 널리 활용되고 있는 Ayoobi et al.[3], FakeNewsNet[4], ISOT[5], Newsroom[6]를 대상으로 한다. FakeNewsNet과 ISOT는 가짜 뉴스를 함께 포함하고 있으며, 이는 추후 Information Blending 생성 기법에 활용된다. 최종적으로 각 데이터셋에서 916 개, 5,362 개, 8,106 개, 15,374 개의 기사가 사용되었으며, 요약문이 없는 뉴스 기사에 대해서는 Gemini-2.0-

flash를 통해 자체적으로 요약문을 생성하였다. 국문 뉴스는 40 만 건의 텍스트 데이터를 포함한 AI Hub 문서요약 텍스트 데이터셋[7]을 활용한다. 이 중 본 연구에서는 10 개 언론사에서 수집된 30 만 건의 뉴스 데이터를 대상으로 하며 원문 평균 길이는 약 1,000 자, 요약문은 150 자로 구성되어 있다. 확보된 원문을 기반으로 AI 생성 뉴스를 생성하며, 요약문은 주제·요지·검증을 위한 비교 기준으로 사용한다.

선행문헌의 LLM-Fake Theory[8]를 기반으로, 다국어 뉴스의 사실성과 문체 다양성을 확보하기 위해 네 가지 AI 뉴스 생성 기법을 설계한다. (1) Sheep's Clothing: 어조를 반전시켜 재작성 함으로써 문체 변이에 따른 다양성을 확보한다. (2) Information Blending: 두 기사의 핵심 요소를 융합해 새로운 사건을 구성하여 사실과 허구의 경계 혼합을 묘사한다. (3) Narrative Generation: 제목, 첫 문장, 요약문을 기반으로 서사를 창조해 뉴스의 구조적 완결성과 자연성을 재현한다. (4) Writing Enhancement: 문체를 객관적·전문적으로 다듬어 사실 보존 하의 편집형 데이터를 생성한다. 또한 AI 뉴스 생성 방법론에 구애 받지 않는 강건한 AI 뉴스 탐지 방법론을 모색하고자, Gemini 2.5 Flash, Gemini 2.5 Pro, Gemma 3 27b, Claude Haiku 3.5, Claude Sonnet 4를 포함한 고성능 대형언어모델(LLM) 5 종을 활용한다. 최종적으로 국문 및 영문 뉴스 9,996 개, 29,758 개를 생성하였다.

### III. AI 생성 및 사람 작성 뉴스 특성 분석 및 결과 검증

\* Equally contributed.

\* Co-corresponding Author.

AI 및 사람 생성 뉴스 간 차이를 관찰하고자 언어적 특성에 대한 정량적 분석을 수행하였다. 영문 및 국문 AI 생성 뉴스는 사람 작성 뉴스에 비해 각 영/한 평균 토큰 수 차이가  $-217.14/-14.52$  로 짧았으며, 평균 문장 길이는 사람 작성 뉴스가  $20.36/16.65$ , AI 생성 뉴스가  $21.34/16.00$  로 유사하나 문장 길이의 표준편차가  $11.86/7.81$ ,  $8.8/6.33$  으로 AI 생성 뉴스에 대한 값이 낮음을 확인했다. 이는 맥락에 따라 다양한 길이의 문장을 작성하는 반면, AI는 균일한 길이의 문장을 구사하려는 경향을 보이는 것을 의미한다.

어휘 다양성 지표인 TTR(Type-Token Ratio)과 MTLD(Measure of Textual Lexical Diversity) 모두에서 영문·국문 TTR  $+0.086$ ,  $+0.042$ , MTLD  $+41.43/+17.0$  로 AI 생성 뉴스가 사람 작성 뉴스에 비해 높은 값을 기록했다. 이는 상대적으로 길이가 짧은 AI 생성 뉴스 내에서 사건·인물·장소 등 다양한 정보를 서술하는 것에서 기인된 결과로 해석된다. 또한 AI 생성 뉴스의 인용문 비율이 영문 및 국문에서 각각  $-0.037$ ,  $-0.057$  로 낮았으며, 이는 AI가 실제 인물 발언이나 인용 구조에 대한 구체적 재현 능력의 한계를 보여준다.

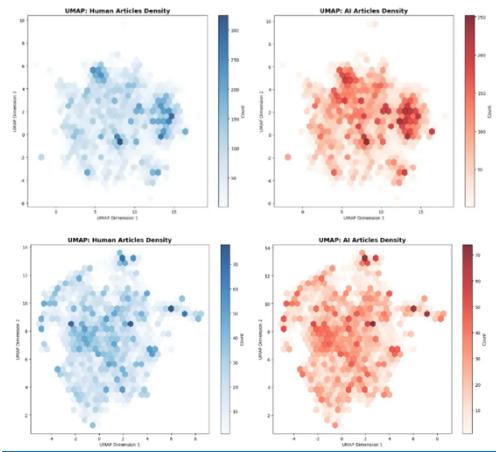


Figure 1 영문(상)·국문(하) AI(빨강)·사람(파랑) 생성 뉴스 UMAP 비교

AI 및 사람 생성 뉴스 간의 차이는 영문 기사에서 보다 두드러지는 경향을 보인다. UMAP (Uniform Manifold Approximation and Projection) 알고리즘으로 기사를 2 차원 공간에 나타낸 결과(그림 1), 영문 기사 대비 한국 기사에 대한 AI·사람 생성 기사 간 차이가 적은 것을 확인했다. 또한 구조적 유사도 비교를 위해 Jaccard similarity( $\uparrow$ )와 Levenshtein distance( $\downarrow$ )를 측정한 결과, 각각 영문·국문에서 평균  $0.355/0.310$ ,  $531.9/2,649.6$  으로, 앞선 분석과 동일하게 영어에 비해 국문 AI·사람 생성 뉴스가 구조적으로 구분이 어려움을 확인할 수 있다.

Table 1 LLM 대상 AI 생성 뉴스 탐지 성능

|    | Accuracy | Recall | Precision | F1-score |
|----|----------|--------|-----------|----------|
| 영문 | 0.830    | 0.660  | 1.000     | 0.795    |
| 국문 | 0.540    | 0.080  | 1.000     | 0.148    |

선행된 분석을 통해 AI 및 사람 생성 뉴스 간 언어적·구조적 차이가 존재함을 확인하였으나, 이러한 차이가 실제 언어모델 수준에서의 탐지 정확도에 영향을 미치는지는 명확하지 않다. 이에 본 연구에서는 클래스 균형을 맞춘 100 개의 뉴스를 대상으로 고성능 LLM 인 GPT-5 를 활용한 AI 생성 뉴스 탐지 분류를 수행하였다.

표 1 과 같이, 영문 대비 국문 탐지 정확도는  $-0.290$  로 대폭 하락했으며, recall 이  $0.080$  으로 대부분의 뉴스를 사람 작성으로 잘못 분류하는 양상이

관찰되었다. 즉 국문일 경우 AI 와 사람 생성 뉴스 간 언어적 변별 신호가 상대적으로 약함을 시사하며, 이는 앞선 분석 결과와도 부합한다.

#### IV. 결론

본 연구는 AI 가 생성한 뉴스와 사람이 작성한 뉴스 간의 언어적 차이를 규명하고자 한다. 이를 위하여 사람이 작성한 영문 및 국문 뉴스 기사를 기반으로 4 가지 AI 뉴스 생성 기법을 적용하여 다국어 AI 생성 뉴스 기사 데이터를 구축했다. 데이터 분석 결과 AI 와 사람 생성 뉴스 간 문체·어휘·구조적 차이가 유의하게 존재했으며, 특히 국문 생성 뉴스는 영문 생성 뉴스에 비해 언어적 변형 폭이 작아 AI 및 사람 생성 여부의 구분이 상대적으로 어려움을 확인하였다. 이러한 분석 결과를 기반으로 LLM 기반 AI 뉴스 시대에서 정보의 신뢰도와 출처를 객관적으로 판단하는 기준 수립에 기여하길 기대한다.

#### ACKNOWLEDGMENT

본 연구는 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받았으며 (No.RS-2025-02217960, 인간의 성장 발달과정을 모사한 인간지향적 AI 모델 개발), 과학기술정보통신부 및 정보통신기획평가원의 학석사연계 ICT 핵심인재양성사업의 연구결과로 수행되었습니다 (IITP-2025-RS-2023-00259497).

#### 참고문헌

- [1] Munoriyarwa, A. & De-Lima-Santos, M.F. (2025). "Generative AI and the Future of News: Examining AI's Agency, Power, and Authority." Journalism Practice, vol. 19, no. 10, pp.2177-2188.
- [2] Terč on, L. & Dobrovoljc, K. (2025). "Linguistic Characteristics of AI-Generated Text: A Survey." arXiv preprint arXiv:2510.05136.
- [3] Ayoobi, N., Shahriar, S. & Mukherjee, A. (2024). "Seeing Through AI's Lens: Enhancing Human Skepticism Towards LLM-Generated Fake News." arXiv preprint arXiv:2406.14012.
- [4] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2019). "FakeNewsNet: A Data Repository with News Content, Social Context and Spatiotemporal Information for Studying Fake News on Social Media." arXiv preprint arXiv:1809.01286.
- [5] Ahmed, H., Traore, I. & Saad, S. (2018). "Detecting opinion spams and fake news using text classification." Security and Privacy, vol. 1, no. 1, e9.
- [6] Grusky, M., Naaman, M., & Artzi, Y. (2018). "NEWSROOM: A Dataset of 1.3 Million Summaries with Diverse Extractive Strategies," Proceedings of NAACL-HLT 2018, pp. 708-719.
- [7] 한국정보화진흥원 AI Hub. (2021). "문서요약 텍스트 (Document Summarization Text)" [데이터셋]. AI Hub (2020/2021 버전).
- [8] Wang, L. Z., Ma, Y., Gao, R., Guo, B., Zhu, H., Fan, W., Lu, Z. & Ng, K. C. (2024). "MegaFake: A Theory-Driven Dataset of Fake News Generated by Large Language Models." arXiv preprint arXiv:2408.11871.