

챗GPT 관련 기사에 대한 토픽 모델링과 분야별 감성 분석

정은희*

Topic Modeling and Field-sepcific Sentiment Analysis for Articles related to ChatGPT

Eun-Hee Jeong*

요 약 챗GPT는 2022년 11월 출시 이후 단 5일만에 100만명의 사용자를 모을 만큼 대중들의 관심이 집중된 대화형 인공지능이다. 본 논문은 2022년 11월 1일부터 2024년 5월 31일까지 빅카인즈에 등록된 챗GPT 관련 기사 중에서 IT, 경제, 국제, 사회, 문화, 정치 분야인 27,796건의 기사를 분석하여 주요 키워드와 11개의 토픽을 도출하고 분야별 감성분석을 하였다. 챗GPT 관련 기사의 주요 키워드는 인공지능, 마이크로소프트, 챗봇, 구글, 반도체, 스타트업 등 IT와 경제관련 키워드의 빈도수가 높게 나왔고, 11개의 토픽은 챗GPT활용과 미래대응교육, 생성형AI 실용화와 주도권 선점, 챗GPT로 인한 관련주 상승, AI반도체 성장, 교수와 학생 학습활동 보조수단, 한국어 인공지능 생태계, 생성형AI 서비스, 챗GPT의 저작권, AI 규제, 오픈AI 개발, 노동시장의 변화이다. 분야별 감성분석에서 IT분야, 경제, 국제, 사회, 문화는 긍정적으로 평가되었으나 정치는 부정적으로 평가되었다.

Abstract ChatGPT is a conversational AI that has attracted the public's attention, gathering 1 million users in just 5 days since its launch in November 2022. This paper analyzed 27,796 articles related to ChatGPT registered on BigKinds from November 1, 2022 to May 31, 2024 in the fields of IT, economy, international affairs, society, culture, and politics, derived major keywords and 11 topics, and performed sentiment analysis by field. The main keywords in articles related to ChatGPT were artificial intelligence, Microsoft, chatbot, Google, semiconductors, startups, etc., and keywords related to IT and economy appeared frequently. The 11 topics are: ChatGPT utilization and future-ready education, practical application of generative AI and taking the lead, rise in related stocks due to ChatGPT, growth of AI semiconductors, support tools for professors and students' learning activities, Korean AI ecosystem, generative AI services, ChatGPT copyright, AI regulations, openAI development, and change of the labor market. In the sentiment analysis by field, IT, economy, international, society, and culture were evaluated positively, but politics was evaluated negatively.

Key Words : Artificial Intelligence, ChatGPT, Sentiment analysis, Text mining, Topic modeling

1. 서론

4차 산업혁명 영역 중 하나인 인공지능(AI : Artificial Intelligence)은 최근 10년 동안 폭발적인 속도로 발전하여, 현재에는 다양한 산업분야에 고도화된 인공지능 기술을 융합하는 디지털 전환이라는 혁신을 주도하고 있다. 특히, 2022년 Gartner는 생성형AI를 기업들에게 새로운 동력을 제공할 것이라 소개하였

다. 생성형AI는 인간과 유사한 창의적 결과물을 생성하는 인공지능으로 주목받고 있을 뿐만 아니라 디자인, 콘텐츠 제작 등 다양한 산업분야에서 새로운 가능성을 보이고 있다.

특히, 챗GPT(Generative Pre-trained Transformer)는 2022년 11월 말 출시 되자마자 대중들의 높은 관심을 받아 단 5일만에 사용자가 100만명이 되었으며, 1개월

만에는 약 1,000만명, 2개월 후에는 월간 활성 사용자가 1억명을 넘어섰다. 이는 다른 IT 기업들과 비교해보면, 9개월인 TikTok, 30개월인 Instagram, 41개월인 Pinterest, 55개월인 Spotify보다 사용자 증가 시간을 크게 단축하였다[1].

챗GPT의 능력이 세상에 존재하는 정보를 합리적으로 조합하여 제공하는 것이라도 그 혁신성에는 이전의 여지가 없다. 챗GPT의 등장은 전 세계에 경제적, 기술적 영향을 미칠 것이 명백하며, 마이크로소프트, 구글 등 빅테크 기업간의 인공지능 기술 패권 경쟁은 이미 시작되었다[2].

챗GPT에 사람들이 열광하는 이유는 첫째, 일반 대중들이 프로그램 학습 없이 접근할 수 있는 대화형 형태로 제공하고, 둘째, 인간의 생산성과 효율성을 높일 수 있으며, 셋째, 다양한 주제와 지식을 포괄하고 사용자의 질문 수준에 따라 맞춤형 답변 제공한다는 것이다[3]. 이렇듯 업계와 일반대중까지 열광하는 챗GPT에 대해 우리 사회는 어떻게 논의하고 반응하는지 탐색해 볼 필요가 있다.

본 논문에서는 사회 여러 분야를 다루고 있는 저널 데이터 중에서 챗GPT에 관련된 IT, 경제, 국제, 사회, 문화, 정치의 뉴스 빅데이터를 이용하여 토픽 모델링을 분석하고, 챗GPT에 대한 분야별 감성은 어떻게 다른지 살펴보고자 한다.

2. 이론적 배경

2.1 챗GPT에 대한 사회적 반응

챗GPT의 등장은 사회 전반에 다양한 변화를 가져왔으며, 사회 각계에서는 챗GPT를 어떤 방식으로 도입하거나 활용해야하는지에 대해 논의가 활발하게 이루어지고 있다.

이호영은 기존 연구 문헌 검토와 무역 전문가 인터뷰를 통해 챗GPT의 무역 업무 활용 가능성을 연구하였으며, 무역업계에서 어떻게 인공지능을 구현할 수 있는지에 대한 이론적, 실무적 시사점을 제시하였다[4]. 강동훈은 챗GPT가 국어교육에 미치는 영향으로 쓰기 교육, 윤리 문제, 거짓 정보 확산 등이 있으며, 이에 대한 대응방안으로 쓰기 교육에 대한 윤리교육 강

화, 비판적 읽기 지도를 통한 비판적 문식성 교육 강화를 제언하였다[5]. 김시현은 텍스트 연결망 분석을 활용하여 챗GPT 관련 언론보도 기사를 분석하여 현 시대는 챗GPT와 같은 생성형AI 도구와 공생하여 살아가는 사회이며 이러한 도구의 활용에 대한 교육적 시사점을 도출하였다[6].

2.2 감성분석

텍스트 마이닝은 비정형 텍스트 데이터 속에 숨어 있는 패턴 또는 의미있는 정보와 지식을 추출해 낼 수 있는 분석 방법으로 빈도, CONCOR, N-gram, TF-IDF, 토픽모델링 및 감성 분석 등 다양한 방법론이 존재한다[7]. 빈도 분석과 토픽모델링 방식은 전통적으로 텍스트 데이터에서 주로 동향이나 경향성을 보여주는데 그치 깊이 있는 논의로 확장하는데에는 제한이 있다.

감성분석은 텍스트에서 감정이나 의견 등을 추출하고 그 의미를 주로 긍정이나 부정으로 파악하는 분석으로 주로 정책 이슈나 논란적인 사건에 대한 SNS, 기사 또는 기사의 댓글 등을 통해 대중들의 여론을 파악하는데 유용하다. 감성분석은 언어적 표현에 들어가는 주관적 요소를 확인하고 이를 수치화하여 텍스트에 나타난 주관성 정도를 평가하는데, 텍스트에서 발견한 주관성 요소를 긍정 또는 부정으로 이분화하여 감정의 방향성을 도출한다. 이는 중립을 포함한 삼분법적 구분, 객관성을 포함하는 사분법적 구분을 통해 주관적 표현이 얼마나 극화되는지 확인하는 방법도 포함한다[8].

기사는 특정 프레임을 내포한 텍스트 형태이며 언론에 의해 조성된 담론은 특히 그 사회와 구성원들이 경험하는 다양한 사건과 현상을 인식하고 해석하는 방향성을 제시한다. 따라서 언론의 중재를 통해 정치, 경제, 과학과 같은 주제들은 대중에게 다양한 형태의 해석 프레임으로 제공된다[9]. 특히, 현대사회에서 언론은 사람들간의 소통과 전달의 핵심 매개체로 작동할 뿐만 아니라 사회 경제적 이슈의 의제 설정의 역할을 한다[10,11,12]. 따라서, 언론기사는 전반적으로 사회, 경제, 정치 등 여러 분야를 다루는 경향이 있기 때문에 연구 데이터로 활용하기에 적절하다.

본 연구에서는 챗GPT에 대한 저널 데이터를 수집하여 토픽 모델링 분석하고 챗GPT에 대한 분야별 감성분석을 하고자 한다.

3. 연구설계

3.1 데이터 수집 및 전처리

본 연구의 목적은 챗GPT에 관련된 IT, 경제, 국제, 사회, 문화, 정치 분야에 대해 이슈를 분석하고, 각 분야별 감성분석을 하는 것이다. 분석 데이터는 빅카인즈에 등록된 2022년 11월 1일부터 2024년 5월 31일 까지 뉴스 빅데이터를 검색어 ((챗GPT) or (ChatGPT) or (챗gpt) or (생성형AI)로 수집하였다. 수집대상 언론사는 전국일간지 12곳, 경제일간지 13곳, 방송사 5곳, 그리고 전문지 10곳으로 총 40곳으로 총 30,019개의 기사를 수집하였다. 수집된 30,019개의 기사 중에서 미분류 869개, 지역 691개, 스포츠 63개, 중복기사 600개를 제외한 IT, 경제, 사회, 정치, 국제, 문화로 분류한 결과는 표 1과 같다. 본 논문에서는 총 27,796개의 기사를 이용해 토픽 모델링과 감성 분석을 하였다.

표 1. 수집된 데이터
Table 1. Collected Data

Category	count	Category	count
IT	14,057	Society	1,333
Economic	8,813	Culture	1,647
International	1,411	Politics	535
Total			27,796

데이터 수집 후에 무의미한 단어, 숫자, 기호 등을 제거하고, 뉴스 빅데이터 검색어인 챗GPT는 모든 토픽에 포함되므로 토픽 모델링과 감성분석에서 제외하고, 인공지능AI, AI, 인공지능 단어들은 인공지능으로 통일하듯이 유사단어를 하나의 단어로 정리하는 전처리 과정을 진행하였다.

3.2 분석 방법

본 연구에서는 토픽 모델링 분석 방법 중에서 가장 일반적으로 많이 활용되는 잠재적 디리클렛 할당 기법인 LDA를 활용하였다. LDA의 토픽은 여러 단어의 혼

합으로 구성되어 있고, 한 토픽에 서로 다른 확률로 여러 단어가 포함되어 있어서, 같은 단어가 서로 다른 확률로 여러 토픽에 포함되어 있다는 가정하에 단어들을 토픽별로 분류하는 기법이다[8].

데이터 분석은 오픈소스 프로그램인 R의 LDA 함수 깁스샘플링을 이용하였으며, 뉴스 빅데이터를 분석하기 위해 하이퍼파라미터 튜닝을 이용하여 적절한 토픽의 수를 선정해 수집회 반복하여 학습하였다. 그 결과 토픽 수는 11개로 선정하였다.

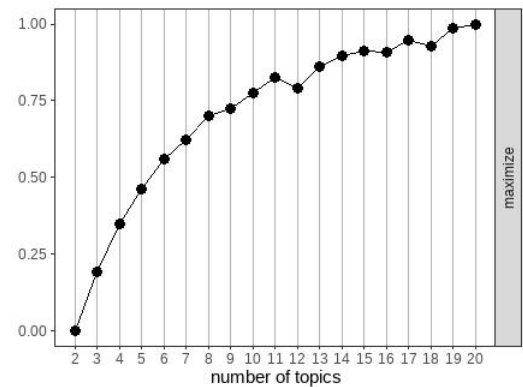


그림 1. 토픽 수
Fig. 1. The number of topic

또한, 군산대학교 감성사전을 이용하여 챗GPT에 대한 감성분석을 실시하였다[13]. 감성사전에서 긍정과 부정의 의미를 동시에 가지고 있는 단어를 부정으로 평가하는 단어들은 중립으로 수정하여 진행하였다. 대표적으로 고도화라는 단어는 상황에 따라 긍정, 부정의 의미를 모두 내포하고 있는데 감성사전에서는 부정으로 평가하고 있어서 중립으로 수정하였다.

4. 분석결과

4.1 키워드 분석

본 논문에서는 IT, 경제, 국제, 사회, 문화, 정치로 구분한 27,796개의 기사에 대한 대표적 키워드 100개에 대한 워드클라우드 분석결과는 그림 2와 같다.

수집된 데이터 중 IT와 경제분야의 비중이 높다보니 빈도수가 높은 키워드들도 인공지능(6,847), 마이

크로소프트(5,788), 이용자(5,046), 한국(4,706), 챗봇(4,466), 구글(4,366), 반도체(4,159), OpenAI(3,801), 스타트업(3,170), 클라우드(2,324), 서울(2,235), 전문가(2,222), 엔비디아(2,179), 샘올트먼(2,123), 중국(2,096), 삼성전자(2,081), 네이버(1,969), 빅테크(1,829), 투자자(1,548) 등처럼 IT와 경제분야로 나타났다.

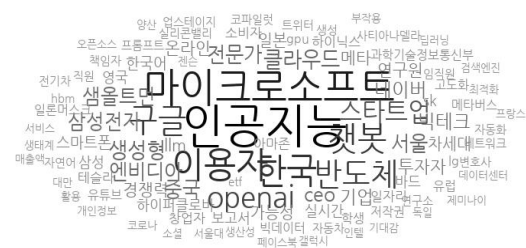


그림 2. 워드클라우드 분석
Fig. 2. Wordcloud analysis

4.2 토픽 모델링 분석 결과

토픽 모델링을 이용해 주요 토픽을 확인할 수 있으며, 토픽 중에서 어떤 토픽이 중요한지는 토픽 비중으로 파악할 수 있다. 본 연구에서는 27,796개의 기사에서 추출된 11개의 토픽과 각 토픽에 대한 비중은 표 2와 같다. 토픽 비중은 생성형AI 실용화와 주도권 선점이 가장 높았으며, AI반도체 성장, 오픈AI 개발, 한국어 AI 생태계, 생성형AI 서비스, 교수와 학생 학습 활동 보조수단, 노동시장의 변화, 챗GPT 저작권, 챗GPT로 인한 관련주 상승, 챗GPT 활용과 미래대응교육 순으로 나타났다.

표 2. 토픽 모델링 결과
Table 2. Topic modeling result

No	Subject	Topic word	Proportion
1	ChatGPT utilization and future-ready education	AI, enterprise, creation, service, utilization, employee, information, model, use, education, blog, human, question, provision, data	0.047

2	Practical application of generative AI and taking the lead	Microsoft, Google, Chatbot, Big Tech, Users, Meta, Bard, Amazon, Copilot, Facebook, Gemini, Search engine, Subscribers, Instagram, Counterpart	0.160
3	Rise of related stocks due to chatGPT	Investor, Researcher, Korea, ETF, Company, Expectation, Sales, Foreigner, Operating Profit, Profitability, KOSDAQ, Related Stock, KOSPI, Net Income, Exchange	0.058
4	Growth of AI semiconductor	Semiconductor, Nvidia, Samsung Electronics, Next Generation, SK, Smartphone, Samsung, Hynix, GPU, Automobile, HBM, Taiwan, Data Center, Intel, Galaxy	0.129
5	Support for professors and students' learning activities	Seoul, expert, online, student, metaverse, Seoul National University, Korea, AI, offline, Ministry of Education, edutech, graduate school, Korea University, research institute, children	0.066
6	Korean AI ecosystem	User, chatbot, Naver, LLM, Korean, Hyperclova, startup, natural language, Upstage, prompt, ecosystem, Naver Cloud, competitiveness, AI	0.113
7	Generative AI services	Generative, cloud, competitiveness, real-time, LG, advanced, network, machine learning, KT, commercialization, consumer, optimization, SKT, customer, differentiation	0.083
8	ChatGPT copyright	AI, possibility, Ministry of Science and ICT, copyright, expert, lawyer, YouTube, media, deep learning, Alpha Go, scientist, research institute, work, influence, KAIST	0.064

9	AI regulation	Korea, China, Japan, UK, Europe, personal information, Germany, EU, side effects, France, Baidu, deepfake, possibility, Russia, fake news	0.092
10	OpenAI development	OpenAI, startup, Sam Altman, CEO, Tesla, founder, Elon Musk, Silicon Valley, Satya Nadella, manager, Twitter, Microsoft, DeepMind, Bill Gates, San Francisco	0.123
11	Change of the labor market	AI, report, jobs, big data, automation, consumers, experts, companies, employees, productivity, online, new technology, staff, competitiveness, efficiency	0.065

11개의 토픽에 대한 주제를 정리해보면,

토픽1은 대화형 인공지능인 챗GPT 활용과 미래 대응 교육이다. 기업, 금융권 등에서는 챗GPT 등장과 행정 활용 주제로 직원 교육을 실시하여 디지털 행정 전환에 대응하고 직원 역량을 강화시켰고, 기업교육 전문기업인 휴넷은 직무 중심 AI 교육과정인 “AI 칼리지”를 출시하는 등 직원 교육이나 업무지원을 위한 인공지능 서비스에 투자하는 것으로 나타났다.

토픽2는 빅테크 기업들의 생성형AI 실용화와 주도권 선점이다. 글로벌 빅테크 기업인 구글은 오픈AI인 챗GPT 대항마로 대화형 생성형AI 챗봇인 제미니를 출시하고, 마이크로소프트는 대화형 인공지능인 코파일럿, 아마존은 맞춤형 고객 비즈니스 서비스를 제공하는 새로운 유형의 생성형AI 비서인 아마존 큐를 출시하는 등 경쟁력을 강화하고, 다양한 분야에서 기술 경쟁을 하고 있다.

토픽3은 챗GPT로 인한 관련주 상승이다. 챗GPT를 중심으로 한 인공지능 열풍으로 주식시장에서는 인공지능에 대한 투자자들의 관심이 증가함에 따라, 인공지능 관련주가 상승세를 나타냈으며, 인공지능 사업에서 성공 가능성을 지닌 플랫폼 기업들이 주목받았다. 그 결과, MS 엔비디아, 인공지능 영상 분석 기술을 주요 사업을 둔 코난테크놀로지, 클라우드 서비스 업체

인 세일즈포즈 등의 기업들의 주가가 상승했으며 증권가는 네이버와 카카오가 데이터와 사용자 확장성을 보유하고 있으므로 인공지능 시장에서 경쟁력이 있는 것으로 전망하고 있다.

토픽 4는 AI반도체 산업의 성장이다. 대용량 학습 데이터를 빠르게 처리하는데 필요한 핵심기술인 AI반도체는 고성능 그래픽처리장치인 GPU보다 효율성이 높아 수요가 증가하는 추세이다. 오픈AI 최고경영자인 샘 올트먼은 한국의 딥테크 스타트업과 AI반도체 역량에 관심이 많으며, AI 반도체를 독점하고 있는 엔비디아에 대응하기 위해 반도체 개발을 함께하고 싶다는 의사를 강력하게 밝혔으며, 삼성과 SK는 파운드리 HBM 협력을 논의 하는 등 글로벌 빅테크 기업들도 적극적으로 AI반도체 개발 및 협력, 투자에 참여하고 있으므로 기술경쟁이 심화될 것으로 보인다.

토픽5는 교수와 학생의 학습활동 보조수단이다. 서울시 교육청이 서울에 있는 교육전문직원 5천여명을 대상으로 챗GPT 관련 설문조사한 결과에 따르면, 챗GPT를 사용해 본 적이 있는 교사는 10명 중 7명으로 조사되었으며, 응답자의 90%가 챗GPT는 교사 역할에 도움이 된다고 응답하였다. 이렇듯, 챗GPT가 교육현장에서 학습자료 조사와 작성, 코딩 생성, 어학학습 등 교수 학습활동의 보조수단으로 활용되고 있으며, 에듀테크 분야로 확대되고 있다. 반면, 시험 부정행위, 논문 대필, 과제 표절 등 악용하는 사례도 늘고 있다. 이러한 인공지능을 이용한 부정행위를 방지하거나 탐색할 수 있는 행동요령이 담긴 챗GPT 종합안내 플랫폼을 개설하거나, 대학의 교수학습 모델에 생성형AI를 접목하여 생성형AI 활용 윤리지침과 수업단계별 AI 활용지침을 마련하는 등 대학들은 학생들이 챗GPT를 올바르게 사용할 수 있도록 노력하고 있으며, AI기술 발전에 맞춰 교육체계를 어떻게 변화시켜야 하는지에 대해 다루고 있다.

토픽6은 한국어 인공지능 생태계이다. 생성형AI 산업이 빠르게 성장하면서 국내 IT 기업들도 한국어에 특화된 인공지능 모델을 출시하고 있다. 카카오는 한국어에 특화된 초거대 인공지능 언어 모델 ‘코GPT’, 네이버는 대화형 인공지능 챗GPT보다 한국어를 6천5백배 더 많이 학습한 초거대 인공지능 하이퍼클로바 X

를 출시하는 등 한국어 특화 초거대 인공지능 시장을 선점하려는 경쟁이 심화될 것으로 보인다. 또한, 네이버, 카카오가 자사 인공지능 기술을 기반으로 다양한 제품 서비스를 개발할 수 있는 스타트업과 손잡고 생태계 확장을 추진하여 더 많은 기업과 고객을 생태계로 유입하려는 전략으로 보인다.

토픽7은 생성형AI 서비스이다. 챗GPT, 하이퍼클로바X 등 기존 생성형AI 서비스는 모두 클라우드를 기반으로 한다. 네이버클라우드에는 AI 최적화 DB지원 기능을 추가함으로써 생성형AI 활용에 대한 기업들의 고민을 줄이고 하이퍼클로바X의 활용도를 높이는 목표를 달성하려고 하며, LG CNS는 마이크로소프트와 손잡고 인공지능, 클라우드 기반 디지털전환 가속화 등 전략적 협업을 통해 신규 서비스를 빠르게 출시하여 시장을 선도하려고 한다. KT는 3대 축인 초거대 AI 상용화, AI 인프라 혁신, AI 인재 양성으로 통신뿐 아니라 모빌리티, 미디어 콘텐츠, 반도체, 기업 공공 디지털 전환 시장을 포괄하는 서비스를 바탕으로 국내뿐 아니라 글로벌 AI 경쟁 시장에 뛰어들었다. 최근에는 AI 클라우드에 의존하지 않으면서 데이터 유출 우려가 상대적으로 적고, 속도가 빠르며 에너지 효율성도 높아 생성형AI 서비스를 이용하거나 개인 맞춤형 서비스를 제공할 수 있는 온디바이스 인공지능 사업에 글로벌 빅테크 기업들의 경쟁이 치열해질 전망이다.

토픽8은 챗GPT의 저작권이다. 2016년 알파고에서 시작된 인공지능이 2022년 말 챗GPT 등장으로 모든 영역에 인공지능이 더해지면서 AI전환시대가 본격화되고 있다. 2023년 인터넷 이용 실태조사에 따르면 국민 절반이 일상에서 인공지능 서비스 사용해 본 경험 있는 것으로 나타났다. 심지어 챗GPT로 작곡을 하거나 소설을 쓰는 등 창작을 하기도 한다. 현재 챗GPT가 만든 창작물의 저작권은 인정하지 않으며, 인공지능 창작물에 사람이 2차 가공한 가공물에 대해 창작자의 저작물로 인정하고 있다. 전문가들은 인공지능 저작권 관련 논쟁은 앞으로 더 거세질 것으로 전망하고 있다.

토픽9는 인공지능 규제법이다. 전 세계적으로 인공지능 기술 경쟁이 확산됨에 따라 개인정보 유출, 딥페이크, 답변의 신뢰성 및 윤리성 등 윤리적인 문제, 기

업 기밀 유출 위험성 등에 대한 우려가 끊임없이 제기되는 가운데, 유럽연합이 처음으로 인공지능을 제한하기 위한 엄격한 규제안인 AI법을 잠정 합의 하였다. AI법은 2026년부터 시행될 예정이며, AI 규제를 위해 고심하는 다른 국가에도 이정표 역할을 할 것으로 보인다.

토픽 10은 오픈AI 개발과 기술 투자이다. 오픈 AI는 샘 올트먼, 그렉 브록만, 일론 머스크, 리드 호프먼 등이 2015년 비영리 회사로 시작하였다. 2018년 GPT 첫 번째 모델을 발표하였으며, 자연어 처리분야에서 새로운 가능성을 제시하여 주목을 받았다. 마이크로소프트는 인공지능 기술개발에 앞서 있던 구글에 대항하기 위해 AI 스타트업인 오픈AI에 2019년 10억 달러를 투자했고, 이후 100억달러 이상 추가 투자를 하였는데, 이는 생성형·대화형 AI 영역에서 구글·애플 등에 밀리고 있는 마이크로소프트가 인공지능 시장을 선점하려는 시도였다. 오픈AI가 2022년 말 생성형 AI 챗봇인 챗GPT를 공개하면서 세계를 인공지능시대로 열었으므로, 마이크로소프트를 인공지능시대에 가장 효과적인 투자를 한 빅테크라 할 수 있겠다.

토픽11은 노동시장의 변화이다. 생성형AI가 고도의 지적 노동뿐만 아니라 예술, 창작에서도 놀라운 역량을 보이고 있으므로 단순 사무직 노동자보다 고임금 지적 노동자를 가장 위협한다고 볼 수 있다. 2023년 구글의 모회사 알파벳은 전직원 6%에 해당하는 12,000명을 감원했고, 마이크로소프트는 게임부문 직원 22,000명 중 9%에 해당하는 약 1,900명을 줄이고, 페이스북도 2,500개의 일자리를 감축하고 신규채용은 하지 않겠다고 밝혔다. 이렇듯 빅테크 기업의 구조조정은 인공지능이 도입된 사무직에 집중되었다. 골드만삭스 보고서에 따르면, 미국과 유럽지역에서 생성형 AI가 일자리의 25%를 대체하고, 세계 노동자의 18%가 일자리를 잃을 것이라 예측하였고, 충격이 가장 클 분야는 행정과 법률분야로 전망하였다. 전문가들은 인간이 인공지능과 일자리를 두고 대결할 것이 아니라 대체할 수 있는 업무는 인공지능에게 시키고 인간은 더 높은 수준의 업무를 수행함으로써 생산성을 향상시켜야 한다고 주장하고 있다. 따라서 인공지능을 잘 사용하는 사람이 더 높은 생산성을 낼 수 있으므로 앞으

로 인공지능 활용능력이 경쟁력이 될 것으로 보인다. 인공지능 발전으로 실직자가 나타나지만 동시에 새로운 직종이 탄생할 것으로 예상되므로 재교육을 통해 더 나은 일자리로 이전할 수 있도록 노동자를 위한 사회안전망이 보장되어야 할 것이다.

4.3 감성 분석 결과

본 논문에서는 수집된 기사를 IT, 경제, 국제, 사회, 문화, 정치로 분류하였고, 분류된 카테고리별 감성분석을 하였다. 각 분야별 감성 분석 결과는 표 3과 같다.

표 3. 감성 분석 결과

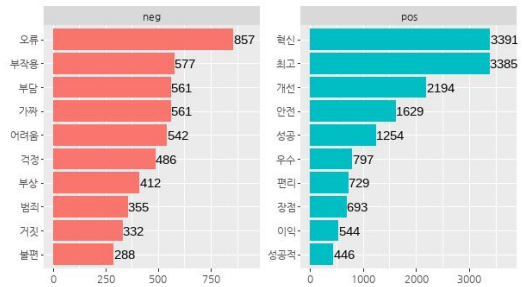
Table 3. Sentiment analysis result

Category	Sentiment	
IT	Negative	10,847
	Neutral	1,980,318
	Positive	18,967
Economic	Negative	6,117
	Neutral	1,243,058
	Positive	12,603
Internaional	Negative	1,689
	Neutral	191,442
	Positive	1,722
Society	Negative	1,481
	Neutral	190,469
	Positive	1,640
Culture	Negative	2,193
	Neutral	242,328
	Positive	2,378
Politics	Negative	957
	Neutral	97,650
	Positive	818

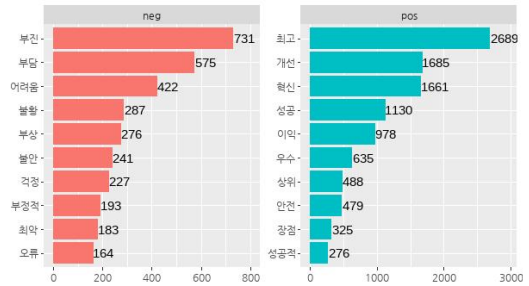
표 3에서 설명하고 있듯이, IT, 경제분야에서는 긍정적인 성향이 두드러지게 나타났고, 사회, 국제분야에서는 큰 격차 없이 긍정적으로 평가되었지만, 정치분야에서는 부정적인 성향으로 나타났다. 정치분야에서는 챗GPT로 대표되는 생성형AI가 거짓 정보 대량 유통, 허위 선전 선동 이미지나 영상물 배포 등 여론 조작에 악용될 수 있다는 우려와 일반 시민 대상 음란물 등 딥페이크 문제가 심각해진 상황 등이 부정적인 성향으로 평가된 것으로 보인다.

각 분야에 대한 부정과 긍정으로 평가하는 단어를 빈도수가 높은 기준으로 살펴보면 그림 3과 같다. 그림 3에서 설명하고 있듯이 각 분야에서 부정과 긍정으로 평가되는 단어들의 유사성이 매우 높다. IT분야에

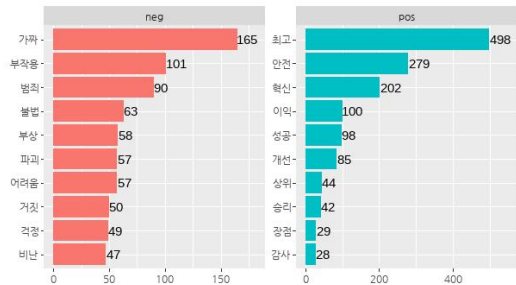
서의 부정단어는 오류(857), 부작용(577) 등이, 긍정 단어는 혁신(3,391), 최고(3,385) 등이 빈도가 높게 나타났다.



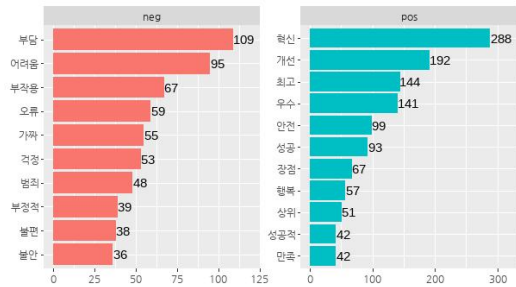
(a) IT



(b) Economic



(c) International



(d) Society

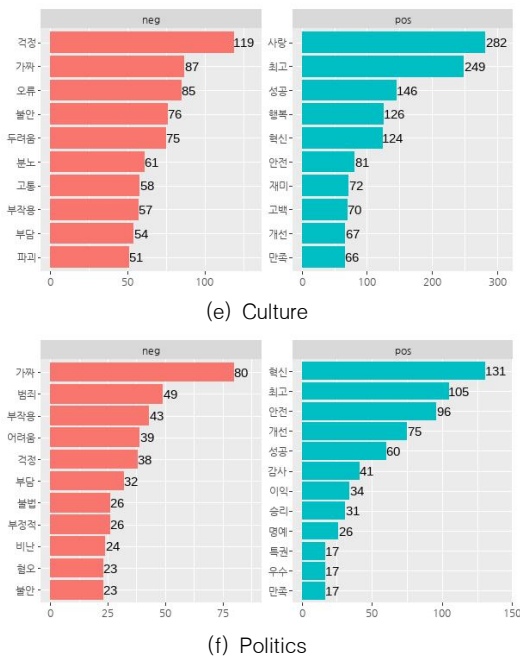


그림 3. 감정 분석 결과

Fig. 3. Sentiment analysis result

경제분야에서의 부정 단어는 부진(731), 부담(575) 등이, 긍정단어는 최고(2,689), 개선(1,685) 등이 빈도가 높게 나타으며, 국제분야에서의 부정단어는 가짜(165), 부작용(101) 등이, 긍정단어는 최고(498), 안전(279) 등이 빈도가 높게 나타났다. 사회분야에서의 부정단어는 부담(109), 어려움(95) 등이, 긍정단어는 혁신(288), 개선(192) 등이 빈도가 높게 나타났고, 문화분야에서의 부정단어는 걱정(119), 가짜(87) 등이, 긍정단어는 사랑(282), 최고(249) 등이 빈도가 높게 나타났다. 그리고, 정치분야에서의 부정단어는 가짜(80), 범죄(49) 등이, 긍정단어는 혁신(131), 최고(105) 등이 빈도가 높게 나타났다.

5. 결론

본 연구는 빅카인즈로 수집한 챗GPT에 관련된 IT, 경제, 국제, 사회, 문화, 정치 분야 기사에 대한 주요 키워드와 토픽을 도출하고, 각 분야별로 감성분석을 하였다.

키워드 분석에서는 인공지능, 마이크로소프트, 챗

봇, 구글, 반도체, OpenAI, 스타트업, 클라우드, 애플, 삼성전자, 네이버, 빅테크 등의 비중이 높은 것으로 나타났으며, 이는 챗GPT 개발에 관련된 빅테크 기업, 인공지능 활용, AI 반도체 등에 관한 기사가 증가한 것을 알 수 있었다.

토픽 모델링의 결과를 종합해보면, 챗GPT 열풍으로 인공지능 관련 기업들의 주가가 상승하고, 생성형 AI 서비스 주도권을 선점하기 위해 AI반도체, AI 클라우드 및 온디바이스 인공지능 등에서 빅테크 기업간의 기술경쟁으로 관련 산업 성장 및 투자 확대 등의 이슈가 두드러졌다. 챗GPT가 교수와 학생의 학습활동의 보조수단으로 도움을 주는 긍정적인 효과와 시험 부정 행위, 논문 대필, 과제 표절 등 부정적인 영향에 대한 양면성 주제가 이슈로 떠올랐다. 기업과 금융권 등에서는 직원교육을 통해 디지털 행정 전환에 대응하기도 하지만, 인공지능의 도입으로 일자리가 축소되어 전세계적으로 노동자의 일자리를 감소할 것이라는 이슈가 도출되었다. 또한, 개인정보유출, 딥페이크, 기업 기밀 유출 위험성, 윤리적인 문제 등을 해결하기 위해 인공지능을 제한하는 AI법과 챗GPT의 창작물에 대한 저작권 관련 이슈가 도출되었다.

감성분석에서는 IT, 경제, 국제, 사회, 문화분야에서는 긍정적으로 평가되었지만, 정치분야에서는 부정적으로 평가되었다. 긍정과 부정으로 평가되는 단어는 유사한 것으로 혁신, 최고, 안전, 우수, 사랑, 개선 등의 긍정적인 단어로 오류, 부진, 가짜, 부담, 걱정, 범죄 등이 부정적인 단어로 나타났다.

1997년 딥블루, 2016년 알파고, 그리고 2022년 챗GPT로 인공지능 기술이 이어져왔다. 초거대언어모델을 기반으로 개발된 챗GPT는 사용자와 단순한 정보를 주고받는 것이 아니라 문학, 미술, 음악 등 다양한 창작 영역에서도 그 진가를 발휘하고 있다. 또한, 챗GPT는 교육, 학습, 투자, 반도체 산업, 노동 시장 등 모든 분야에 영향력을 미치고 있다. 따라서 인공지능 사용에 대한 보안 인식과 사회적 윤리성 강화가 필요하며, 챗GPT의 활용으로 인한 실업문제에 대한 대응으로 직업교육이나 사회정책적 측면의 보장에 대한 논의가 필요하다.

REFERENCES

- [1] D. S. Go, "ChatGPT, Monthly Active Users Surpass 100 Million", Available : <https://www.cwn.kr/news/articleView.html?idxno=15060>
- [2] E. Y. Lee, S. J. Oh, S. E. Kang, H. W. Choi and S. Y. Shin, "ChatGPT, Opportunity or Threat?", PwC Korea, 2023.
- [3] T. J. Kim and S. L. Han, "An Exploratory Study on social Issues Related to ChatGPT : Focusing on News Big Data-based Topic Modeling Analysis", Journal of Digital Contents Society, vol.24, no.6, pp.1209-1220, Jun. 2023.
- [4] H. Y. Lee, "Exploratory Study on the Utilization of ChatGPT in International Trade", International Commerce and Information Review, vol.52, no.1, pp.43-61, Mar. 2023.
- [5] D. H. Kang, "The advent of ChatGPT and the response of Korean language education", Korean Language and Literature, vol.82, pp.469-496, Mar. 2023.
- [6] S. H. Kim and S. H. Yang, "Analysis of Domestic Press Coverage on ChatGPT using Text Network Analysis", Journal of Educational Innovation Research, vol.33, no.4, pp.303-328, Dec. 2023.
- [7] Y. W. Kim, "Do it! R Text Mining," Easypublishing, Jan. 2021.
- [8] C. S. Choi and Y. H. Im, "Sentiment analysis as a research method on partisanship in the presidential coverage", Korean Journal of Journalism & Communication Studies, vol.65, no.1, pp.35-70, Jan. 2021.
- [9] H. N. Kwon, S. Zhang and C. J. Chung, "Analysis of Changes in AI Perception Using Text Mining from AlphaGo to Chat GPT by Focusing on Legacy Media and Social Media", Discourse and Policy in Social Science, vol.16, no.2, pp.205-240, Oct. 2023.
- [10] S. S. Lee, "Analysis of Subject Category on Artificial Intelligence Discourse in Newspaper Articles", Korean Library And Information Science Society, vol.48, no.4, pp.21-47, Dec. 2017.
- [11] H. J. Lee, C. S. Sung and B. H. Jeon, "GenAI(Generative Artificial Intelligence) Technology Trend Analysis Using Bigdata: ChatGPT Emergence and Startup Impact Assessment", Asia-Pacific Journal of Business Venturing and Entrepreneurship (APJBVE), vol.18, no.4, pp.65-76, Aug. 2023.
- [12] G. J. Yee, "The Media Crisis and Democracy in Korea", Korea and Global Affairs, vol.7, no.3, pp.537-564, Jun. 2023.
- [13] Kunsan National University KNU Korean Emotion Dictionary, Available : <http://github.com/park1200656/KnuSentiLex>

저자약력

정 은 희 (Eun-Hee Jeong)

[중심회원]



- 1998년 2월 : 관동대학교 일반대학원 전자계산공학과(공학석사)
- 2003년 2월 : 관동대학교 일반대학원 전자계산공학과(공학박사)
- 2003년 9월 ~ 현재 : 강원대학교 글로벌인재학부 경제금융전공 교수

〈관심분야〉 빅데이터 분석, 텍스트 마이닝, 계량경제, 경제통계