<저널리즘의 이해 보고서>

2018240017 배한민 파트

<서론>

텍스트, 전자제품, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명위 기사는 챗 GPT가 작성한 기사입니다. ‘챗 GPT가 언론에 미치는 영향’을 기사로 써달라고 하면 몇 초만에 이런 기사를 만들어 냅니다.

인공지능이 발전함에 따라 저널리즘이 위협받고 있습니다. 챗 GPT가 사용되기 이전부터 인공 지능이 뉴스에서 어떻게 사용되어 왔는지를 알아야 챗 GPT에 대응할 수 있을 것입니다. 또 그 후 저널리즘에서 Chat GPT가 어떻게 사용되어야 하는지, 이 사회 속에서 기자들은 어떻게 나아가야하는지는 꼭 탐구가 필요한 문제입니다. 그래서 저희 조는 이 주제를 탐구하기로 했습니다.

<인공지능의 정의, AI 뉴스 서비스의 정의>

인공지능이란 인간의 지능이 가지는 학습, 추리, 적응, 논증 따위의 기능을 갖춘 컴퓨터 시스템을 말합니다. 이런 인공지능이 쓰인 뉴스를 말하는 ‘AI 뉴스 서비스’는 구체적으로 취재, 기사 작성, 콘텐츠 유통 등 언론사 업무 흐름 중 특정 과업을 AI에 맡기는 체계를 말합니다.

ex. ai가 기사를 분석해 해당 내용에 어울리는 사진을 자동으로 추천해주는 서비스나 ai가 기상 데이터를 가져와 속보 기사의 초안을 쓰는 서비스 등 도 예시가 될 수 있습니다.

<뉴스에 관여하는 ai의 종류>

뉴스에 관여하는 ai의 종류는 먼저 두 가지로 나눌 수 있습니다.

1. 첫번째로 룰 기반 ai가 있습니다.

1950년대부터 지금까지 쓰여온 터줏대감 ai로 ‘A 라는 데이터가 들어오면 B라는 결과값으로 바꿔라’라는 명령, 룰의 원칙을 사용하는 ai입니다.

ex. 로봇 기사 서비스

* 룰기반 ai는 안정적이지만 간단한 일만 할 수 있습니다.

1. 두 번째로 머신 러닝 기반 ai가 있습니다.

머신 러닝은 ‘기계학습’이라는 말뜻대로 기계가 스스로 특정 능력을 배우고 키워가게 해주는 기술입니다. 기계는 대량의 데이터를 공부해 특정한 패턴을 계속 찾아내며 실력을 쌓죠.

이 머신러닝 ai는 복잡한 일을 유연하게 처리 가능. 좀 더 우리가 요즘 잘 아는 챗 gpt와 비슷한 종류의 인공 지능이죠.

ex. 사람의 말을 듣고 이해하는 스피커, 이미지에서 특정 색깔을 찾아주는 ai 엔진

* 머신러닝 기반 ai는 더 똑똑한 일을 할 수 있는 만큼 가격이 비쌉니다.

저희는 룰 기반부터 머신러닝까지 어떻게 인공지능이 지금의 시대까지 오게 되었는지 인공지능 AI 뉴스 서비스의 역사를 먼저 알아보도록 하겠습니다.

<AI 뉴스 서비스의 역사 (해외 편)>

1. 톰슨의 자동 기사

다양한 분야에서 ai가 쓰일 수 있지만 첫 AI뉴스 서비스는 자동 기사(로봇기사)였습니다. 캐나다의 금융정보회사 톰슨이 2006년 3월 도입. 기업 실적 데이터를 수신해 0.3초만에 기사를 만들어냈습니다. 이는 톰슨이 금융 관련 구조화 데이터를 많이 확보하고 있었기 때문에 가능했던 일이죠.

1. LA타임스의 사건 사고 특화 로봇기사

다음으로는 미국 <LA 타임스>가 사건 사고에 특화한 로봇 기사 서비스를 선보였습니다.

2010년 LA시를 포함하는 광역 행정구역인 LA 카운티의 검시 자료를 분석해 자동으로 기사를 쓰는 로봇을 만들어냈습니다. (지역의 살인 범죄 양상을 조망하는 것이 목적이었음) 살언사건 데이터에는 사건이 발생한 일시와 장소, 피해자의 인종과 나이 등이 포함돼 있는데, 로봇은 이를 토대로 다음과 같은 문장을 생성했습니다. ‘라틴계인 트래비스 솔로리오(26)씨가 5일 선밸리 태머랙가 9300 구역에서 총격을 당해 숨졌다고 LA 카운티 검시관이 밝혔다.” 이렇게 로봇이 리드 문장을 만들어 보내면 여기에 이제 인간 기자가 취재한 내용을 더해 기사를 완성해나가는 식입니다. 로봇은 사건 발생 지역의 지도도 기사에 첨부할 수 있습니다.

1. LA 타임스의 퀘이크봇

LA 타임스는 2011년 AI 지진 기사 서비스도 선보였습니다. ’퀘이크봇’은 미국 지질조사국의 지진 정보를 받아 진앙과의 거리와 지진의 규모를 기준으로 기사 작성 여부를 판단, 3분만에 바로 기사를 만들어냅니다. 이 퀘이크 봇의 등장은 정말 혁신적이었습니다. 덕분에 기자들은 지진에 대한 정보를 실시간으로 확인하고 뉴스를 작성할 필요 없이 퀘이크봇이 작성한 기사를 검토하고 발행하는 정도만 하면서 단순 기사에 드는 시간낭비를 줄일 수 있게 되었습니다.

인공지능은 사람과 다르게 실수를 할 일은 없지만, 데이터가 잘못 입력되면 오보를 하기도 합니다.

→ 퀘이크 봇은 2017년 6월 22일 있지도 않은 지진 속보를 냈습니다. 샌타바버라에서 약 16km 떨어진 태평양 연안에서 규모 6.8의 지진이 발생했다는 오보를 했습니다. 원인은 미국 지질 조사국의 실수였습니다. 한 연구자가 100년 전에 발생했던 지진 데이터를 다루다가 실수로 경보 시스템을 건드렸는데, 이 데이터가 퀘이크 봇에 입력된 것입니다. 이후에 더 보안을 강력하게 해서 추가로 오보는 없었다고 합니다.

1. AP통신의 기업 실적 기사

다음으로 <AP통신>이 2014년 기술 스타트업 ‘오토메이티드 인사이츠’와 제휴해 AI가 매출, 영업, 손익, 순손익 등 기업 실적 기사를 쓰게 했습니다. AP는 이 AI 기사 덕분에 기자들의 업무 시간은 20% 줄어든 반면, 기업 실적 기사의 건수는 12배이상 늘었습니다. 이후에도 해외에서는 계속 AI를 발전시키며 저널리즘에 이용해왔습니다.

<현재 우리가 주변에서서 볼 수 있는 AI 뉴스 서비스들>

취재와 보도로 이루어지는 저널리즘의 행위는 세부적으로는 매우 다양하기에 현재까지 AI 알고리즘이 모든 저널리즘 행위를 대체하지는 못하고 있습니다. 하지만 뉴스의 가치를 평가하고 이를 토대로 뉴스 노출의 우선순위를 결정하는 등 저널리즘 전반에 걸쳐서 각각의 부분에서 AI 알고리즘의 활용은 확대되는 추세를 보이고 있습니다. 취재 시작부터 뉴스라는 결과물을 내놓고 이를 뉴스 소비자에게 노출하거나 전달하는 국면에서 AI 알고리즘의 개입 예시를 다음과 같이 분류할 수 있습니다.

1. 기사 작성, 요약, 뉴스 작성의 자동화
2. 연합 뉴스에서 기사 3줄 요약 기능 인공지능으로 탑재한 것

연합뉴스에서는 기사 뷰페이지 상단에 본문 대용을 3줄로 요약하는 ‘기사 본문 ai 3 줄 요약’ 서비스를 오픈했습니다. 다음은 이 서비스의 개발 과정입니다.

먼저 알고리즘과 뉴스 요약용 학습 도구를 개발하고, 주제별 요약 데이터를 활용해 기사 요약 방법을 이식했습니다. 그 후, 사람이 요약한 자료를 알고리즘에 사전에 학습시켜 AI 스스로 긴 기사를 짧은 문장으로 요약할 수 있도록 만들었습니다. AI가 요약한 결과물은 점수로 환산해 사람이 다시 재평가하고 이 과정을 통해 AI 학습 알고리즘을 정교화할 수 있었습니다.

1. 기사 작성뿐만 아니라, 방송사에서는 AI 앵커로 사람들의 주목도를 높이고 있습니다. MBN은 김주하 AI앵커, YTN은 변상욱 AI 앵커를 선보여 뉴스를 진행했습니다.

1. 뉴스의 노출 순위 결정

취재 과정을 거쳐서 생산한 뉴스는 최종적으로 해당 뉴스에 대한 가치를 평가하고, 노출 순위를 정해서 소비자에게 전달됩니다. 보통 이 과정은 뉴스 생산의 마지막 국면에 해당하는데 전통적으로 편집 기자들의 전문적이고 고유한 업무로 간주되어 왔습니다. 이 과정에서 편집 기자들은 객관적인 관점에서 기사의 중요도, 독자를 포함한 일반인들의 관심도, 가독성 여기에 자사의 편집 방침을 감안해서 뉴스의 우선순위를 결정해왔습니다. 자사의 편집 방침은 언론사마다 고유의 원칙과 철학을 말하는데 이에 따라서 뉴스의 노출 순위는 종종 큰 차이를 보이게 됩니다.

이렇듯 전통적으로 편집 기자들이 수행하던 뉴스의 우선순위 결정에 AI의 개입이 증가하고 있습니다. 뉴스의 양이 급격하게 늘어나고 그만큼 이용자들의 요구도 다양해지는 가운데 수많은 뉴스들을 빠르고 상품성 있게 배열하고 뉴스 배열에서 필연적으로 발생하는 객관성 논란을 피하기 위한 목적으로 AI를 사용합니다.

최근 들어 언론사를 비롯해 뉴스 플랫폼들이 보편적으로 활용하고 있는 뉴스 랭킹 알고리즘이 대표적인 사례입니다. 뉴스 랭킹 알고리즘은 획일적이라는 단점이 있는데 이런 문제점을 보완해 최근 들어 가장 많이 사용되는 뉴스 배열 알고리즘이 ‘맞춤형 추천 알고리즘’입니다. 뉴스를 소비하는 개인들의 성향이 다른 만큼 개인 수요자가 가장 읽고 싶거나 보고 싶을 것으로 추정되는 뉴스 기사를 서열화해서 추천하는 방식이죠.

1. 뉴스의 객관적 가치 평가

뉴스 클러스터링(clustering) 알고리즘은 뉴스 가치를 판단해 주는 가장 대표적인 알고리즘입니다. 클러스터링은 수집된 뉴스들로부터 공통된 주제를 찾는 것과 정해진 주제에 따라 유사한 뉴스들을 묶는 것 모두를 말합니다. 실제로 네이버, 다음 등 포털에는 하루 3만 개 이상의 뉴스가 유입되고 있는데 AI 알고리즘은 이 가치 판단 행위를 일정한 원칙에 따라 자동화하여 사람의 가치 판단 행위를 대체하고 있습니다

텍스트, 스크린샷, 웹사이트, 웹 페이지이(가) 표시된 사진

자동 생성된 설명오른쪽 사진은 2019년 5월 27일 오전 10시, 검색어로 ｢검찰개혁｣을 입력했을 때 네이버와 다음이 제시하는 클러스터입니다. 두 검색엔진은 검찰개혁이라는 검색 키워드와 일치하는 기사들을 우선적으로 제시한다는 공통점을 확인할 수 있다. 특히 첫 번째 클러스터의 경우 언론사 배치는 상이했지만, 특정 인물과 관련된 검찰개혁 기사를 똑같이 배치한 것을 통해서 두 포털은 클러스터링의 기본 원칙을 공유하거나 해당 클러스터링에 대해 서 비슷한 관점에서 평가하는 것으로 유추해 볼 수 있습니다. 이에 비해서 두 번째 클러스터부터는 두 포털의 클러스터링 기법이 차이를 보이고 있습니다. 다음의 경우 검색화면 전체가 동일한 인물과 관련한 내용의 기사를 제시하고 있지만, 네이버의 경우 두 번째와 클러스터에는 다른 인물을 언급한 기사를 배치하고 있습니다. 특히, 네이버는 검색어, 검찰개혁과 관련해 첫 화면에 제시한 네 개의 클러스터 가운데 첫 번째와 세 번째, 두 번째와 네 번째 인물을 중복하면서 다음과의 차별성을 도모하고 있는 것으로 보입니다. 이를 비추어 볼 때 네이버와 다음은 기본적인 알고리즘은 유사한 클러스터링 기법을 채용하고 있지만, 세부 항목에서는 독자적으로 개발한 알고리즘이 작용하고 있고, 이로 인해서 세부적인 항목으로 다운사이징 할수록 두 포털의 검색결과는 상이하게 달라지는 것을 확인할 수 있습니다

<레퍼런스>

이상도. "로봇 저널리즘 등장에 따른 한국 언론의 변화와 발전방향에 관한 연구." 국내석사학위논문 동국대학교, 2017. 서울

오세욱.(2016).저널리즘과 알고리즘의 융합에 대한 탐색적 연구.사이버커뮤니케이션학보,33(3),51-101.

김태균 권영전 박주현, AI 뉴스 제작 방법, 2021, 커뮤니케이션북스

이윤영. "AI 저널리즘 활성화를 위한 전문가 집단의 인식 및 사례 연구." 국내박사학위논문 청주대학교, 2019. 충청북도

서명덕, 2021.01.11, 연합뉴스, 인공지능 기사요약 서비스 첫선, 연합뉴스, <http://www.mediatoday.co.kr/news/articleView.html?idxno=308243>

금준경, 2023.02.01, ‘챗GPT’는 기자에게 어떤 영향을 미칠까, 미디어오늘, <http://www.mediatoday.co.kr/news/articleView.html?idxno=308243>