

[SPRI AI BRIEF]

인공지능 최신 동향과 시사점



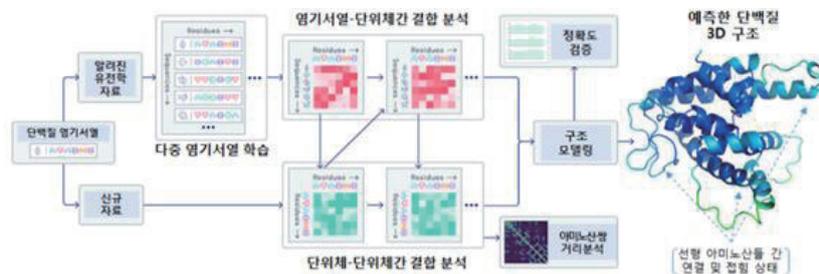
CONTENTS

- 팀마인드, 시로 50년 된 생명과학분야 난제 해결
- 서비스나우, 엘리먼트AI 인수 발표
- 유튜브, 시로 유해 동영상에 이어 악성 댓글 관리 강화
- 구글, AI 학습 방식의 한계에 주목

딥마인드, 시로 50년 된 생명과학분야 난제 해결¹

- 구글 딥마인드(DeepMind)사는 50년간 풀지 못한 단백질 접힘(Protein Folding) 문제를 해결한 AI, 알파폴드(AlphaFold)를 개발하여 공개
 - 생명체 기본 구성요소인 단백질*은 선형 아미노산(Amino-acid)복합체의 염기서열에 따라 각각 고유한 접힘(Folding) 구조를 가짐
 - * 현재 알려진 단백질은 약 2억 개이며, 매년 3천만 개가 신규로 발견
 - 인체의 모든 세포, 조직, 기관은 아미노산이 20가지 형태의 독특한 순서와 조합으로 만들어진 10만 가지 이상의 단백질로 구성되어 특정 기능을 수행
 - 지난 50년간 연구자들은 단백질의 3차원 구조를 예측하고자 했으나, 가능한 배열의 수가 천문학적 (10^{300})이기에 효과적 예측이 불가
- 알파폴드는 딥 뉴럴 네트워크(DNN)를 활용하여 방대한 게놈(Genome) 데이터를 학습하였고, 아미노산 서열에 기반 하여 단백질의 3차원 구조를 예측
 - 약 17만 개의 아미노산 서열과 단백질의 모양을 학습하여, 아미노산 쌍 사이의 거리와 아미노산들을 연결하는 화학결합 각도를 기반으로 단백질 구조를 예측

[그림 1] 알파폴드 학습모델 구조



※출처 : 딥마인드, 재편집

- 연구진은 2016년부터 단백질 접힘 예측을 위한 연구개발을 추진해왔으며, 세계 단백질 구조 예측 대회 (CASP)*에서 우승하며 알파폴드의 성능을 입증
- * 1994년부터 격년으로 개최되는 연구 대회로 미지의 단백질 100개에 대한 아미노산 서열을 주고 이를 통해 구조를 예측, 딥마인드는 지난 2018년에 이어 올해도 92.4점으로 우승
- 생명과학 분야의 AI 활약으로 새로운 질병을 예측하고 난치병 치료법을 개발하며 의료 및 제약 산업에 매우 큰 파급효과를 가져올 것으로 기대
 - 단백질의 접힘을 결정짓는 원리를 이해하면, 아미노산 서열로부터 특정 단백질의 고유한 구조를 예측할 수 있고, 단백질의 역할 및 생명체의 작동 원리 파악도 가능
 - 인체 내의 잘못 접힌 단백질 구조는 알츠하이머, 파킨슨, 당뇨 등 난치병을 유발하며, 단백질을 분석하여 감염병 치료방법 및 변종 등을 예측*하는 것이 가능
 - * 딥마인드는 알파폴드를 이용해 신규 코로나-19 변종을 6개 발견

¹ The Guardian, "DeepMind AI cracks 50-year-old problem of protein folding", 2020.11.30. DeepMind blog, "AlphaFold: a solution to a 50-year-old grand challenge in biology", 2020.11.30.

서비스나우, 엘리먼트AI 인수 발표²

● 미국의 클라우드 기반 IT서비스 기업인 서비스나우(ServiceNow)*는 캐나다의 AI선도 기업인 엘리먼트AI(Element.ai)社를 인수한다고 발표(11.30.)

* 2004년 설립되어 전 세계 6,300개 이상의 기업 고객, 34.6억 달러 매출(2019), 직원 12,500여명(2020)을 보유한 글로벌 IT서비스 관리 전문 기업으로 현재 시가총액은 약 1,000억 달러

- 서비스나우는 자사의 IT서비스 관리 플랫폼 자동화, 지능화 기술 역량 확보에 적극 투자 중이며 올해 들어 엘리먼트AI 포함, 4개의 AI기업을 인수*

* Passage AI(대화형 AI기술, '20.1), Loom(시스템 장애예측, '20.1), Sweagle(데이터관리, '20.6)을 각각 3천 3백만 달러, 5천 8백만 달러, 2천 5백만 달러에 인수

- 엘리먼트AI 인수 금액은 약 5억 달러*로 예상되며 2021년 상반기 완료 예상

* 엘리먼트AI는 마이크로소프트, 인텔, 엔비디아, 텐센트 등 유수 기업으로부터 약 2억 5천만 달러 이상 투자받았으며 2019년 9월 추가 투자 유치 시 약 6~7억 달러의 가치 평가를 받았음

● 인수를 통해 AI기술 및 전문 인력을 확보하여 자사 플랫폼의 AI 역량 강화 추진

- **(AI인력)** 서비스나우는 엘리먼트AI의 기술 전문 인력을 통해 자사 워크플로우 관리 플랫폼(Now Platform) 고도화에 집중하고 엘리먼트AI의 기존 고객 서비스는 중단할 계획

- 엘리먼트AI는 최고 수준의 박사급 인력 100여명을 포함해 500여명의 인력 보유

- **(혁신허브)** 엘리먼트AI가 소재한 몬트리올에 AI Innovation Hub를 설치해 엘리먼트AI의 고급 인력을 중심으로 AI기반 플랫폼 기술 개발 계획

- 엘리먼트AI의 공동창업자이자 세계적 석학인 몬트리올 대학 요슈아 벤지오 교수가 인수 후에도 서비스나우의 기술 고문으로 참여할 전망

● 팬데믹 시대 디지털 전환 관련 사업기회 선점을 위한 IT기업 간 합종연횡 가속화 전망

- 코로나-19의 장기화로 업무 자동화, 재택 근무 확대, 비대면 협업 등 일하는 방식이 변화하고 있으며 이를 지원하기 위한 IT서비스의 지능화 필요성 증대

- IT서비스 제공 기업 간 기업의 운영 프로세스 자동화 및 통합적 관리를 지원하는 최적의 솔루션 확보 경쟁 심화*

* 서비스나우 한국 진출 및 LG CNS와 파트너십 체결 (2019.11), 세일즈포스의 기업용 메신저 업체 슬랙(Slack) 인수 (약 30조원) (2020.12), 마이크로소프트 협업 플랫폼 팀즈 서비스 출시(2017), 네이버(라인웍스)·카카오엔터프라이즈 등 국내 기업용 통합 플랫폼 시장 진출

- 기술력과 인재를 보유했으나 비즈니스 모델이 부재하여 매출이 적은 AI 스타트업의 인수 합병이 늘어날 전망*

* 2016~2019년까지 AI기업의 인수합병은 87건에서 231건으로 2배 이상 증가(CB Insight)

- 이번 엘리먼트AI사의 피인수도 뚜렷한 사업 모델이 없는 유망 AI스타트업이 거대 테크 기업들과의 경쟁에 한계를 느낀대서 비롯됐다는 견해도 존재

² ServiceNow, "ServiceNow to Acquire AI Pioneer Element AI", 2020. 11. 30.

유튜브, 시로 유해 동영상에 이어 악성 댓글 관리 강화³

- 유튜브는 2020년 12월 3일 시를 이용해 부적절한 댓글에 알림 팝업을 띄우는 새로운 커뮤니티 가이드라인을 시행한다고 발표
 - 작성자에게 현재 포스팅하려는 댓글이 적절한지 가이드라인을 읽어보라고 권고하고 그대로 게시하거나 수정할 수 있도록 선택권 부여
- 유튜브는 그동안 시를 활용해 유해 동영상 관리를 해왔으며, 이번 악성 댓글 차단에도 시를 적극 활용하기로 결정
 - 유튜브는 유해 동영상을 시가 판단하고 관리자가 재확인하는 방식을 통해 관리⁴
 - 급증하는 유해 동영상을 처리하기 위해 시를 적극 활용하기로 결정한 이후, 2020년 3분기에 유튜브가 삭제한 동영상은 약 787만 여개
 - 같은 기간에 180만 개 이상의 채널이 커뮤니티 가이드를 3회 이상 위반하거나 심각한 악용 등으로 폐쇄 조치
 - 증오 표현, 잘못된 정보, 괴롭힘 등 유해 여부 판단이 애매한 경우가 많아 시가 부적절하다고 판단한 동영상을 인간이 재확인 후 최종 결정
 - 유튜브는 최근 부적절한 내용이 포함되어 삭제된 댓글이 2019년 초 대비 46배나 늘어 시를 이용한 관리 강화를 결정⁵
 - 이용자들이 반복적으로 신고한 내용을 시가 학습하고 공격성이 있다고 판단되는 댓글을 식별
 - 학습이 충분치 않아 오류가 있을 수 있음을 인지한 유튜브는 도입 초기 단계임을 고려해 댓글 작성자의 참여*를 유도해 시 성능을 테스트
 - * 작성자 스스로 공격적이지 않다고 판단되는 댓글에 알림 팝업이 뜬 경우, 알림(Let us know) 버튼을 클릭해 유튜브 측에 피드백이 가능
- 인터넷 댓글 관련 문제가 끊이지 않는 가운데 악성 댓글을 감지·차단할 수 있는 자연어 처리 기술의 고도화가 유망 활용 분야로 부상
 - 네이버는 올해 6월부터 댓글에 욕설이 없더라도 문맥상 모욕적이거나 타인을 비하하는 표현이 있을 경우 악플로 인식하는 '클린봇 2.0*'을 출시
 - * 악플 적발 정확도를 95%로 판단하고 있으나 현재 시 성능으로는 여러 변칙적 악성댓글 대응에 취약해 올해 3월 연예 뉴스에 이어 9월 스포츠 뉴스 댓글 기능을 폐지하기로 결정

³ Youtube, "Help us keep comments respectful - new Community Guidelines comment reminders", 2020.11.9.

⁴ 구글 투명성보고서, "YouTube 커뮤니티 가이드 시정 조치"

⁵ Cnet, "YouTube will ask you to rethink posting that comment if AI thinks it's offensive", 2020.12.3.

구글, AI 학습 방식의 한계에 주목⁶

- 현재 AI는 연구개발 단계에서의 탁월한 성능이 실제 활용 단계에서 급격히 떨어지는, 소위 '데이터 시프트(Data Shift)' 문제가 빈번히 발생
 - AI 혁신을 주도하는 심층학습(Deep Learning)은 실제 데이터가 학습한 데이터와 형식과 속성이 다른 상황, 즉 데이터 시프트가 발생한 경우 성능 저하
 - 가령 고해상도의 이미지 데이터로 학습한 AI는 저해상도 이미지나 이미지 조각을 인식해야 하는 실제 상황에서 제대로 작동 안함
- 최근 구글 연구진은 현실세계에서 AI의 성능이 하락하는 또 다른 원인으로 비명세성(Underspecification) 문제를 제시⁷
 - 구글 연구진은 연구개발 단계에서 성능이 유사한 것으로 판명된 50종의 이미지 인식 AI 모델이 실제 상황에서 매우 다른 성능을 발휘함을 발견
 - 50종의 AI 모델은 무작위로 부여한 초기 값만 다를 뿐 동일한 '이미지넷 데이터'를 이용해 학습했으며, 테스트 결과 성능이 유사함을 확인
 - 이후 이미지넷 데이터의 명암, 대비를 조정한 '이미지넷-C 데이터'와 객체의 위치를 조정한 '오브젝트 넷'을 활용해 실제 상황에서의 성능을 테스트
 - 테스트 결과, '이미지넷-C 데이터'를 잘 인식하는 모델과 '오브젝트 넷'을 잘 인식하는 모델이 구분되는 등 실제 상황에서 모델 간 성능 편차가 상당히 존재함을 확인
 - 비명세성은 동일한 현상을 예측하는 무수히 많은 통계모델이 있듯이, AI도 학습 과정에서 이미지를 인식하는 방식이 미묘하게 다른 다수의 모델이 탄생함을 의미
 - 가령, 경제전망기관은 경제성장률을 결정하는 무수한 요인 중에서 가장 중요하다고 판단하는 변수를 선택하여 고유한 전망모델을 구축하며, 경제 상황별 기관 간 정확도가 상이
 - AI 역시 초기 값 설정, 학습 횟수, 한 번에 학습시킬 데이터의 양 등 사소한 차이가 인식 방식이 상이한 모델을 만들고, 현실 세계에서 상황에 따른 성능 편차가 발생
- AI 모델의 비명세성은 AI의 성공적인 산업적 활용을 위해 극복해야 할 과제
 - 이번 구글 논문의 성과는 현실세계에서 불안정한 AI 성능을 비명세성으로 설명했다는 점에서, AI의 활용 확대에 대한 단초를 제시함
 - 하지만 비명세성의 해결 방안은 아직 없으며, 실험실 테스트를 통과한 AI 모델 간 실제 성능 편차를 줄이기 위한 추가적인 개발 과정을 도입해야 할 필요성을 시사

⁶ MIT Technology Review, "The way we train AI is fundamentally flawed", 2020.11.18.

⁷ Google, "Underspecification Presents Challenges for Credibility in Modern Machine Learning" (2020.11.)



홈페이지 : <https://spri.kr/>

* 보고서와 관련된 문의는 시정책연구팀(wycho@spri.kr, 031-739-7312)으로 연락주시기 바랍니다.