

KAKAO

Vol.12

AI

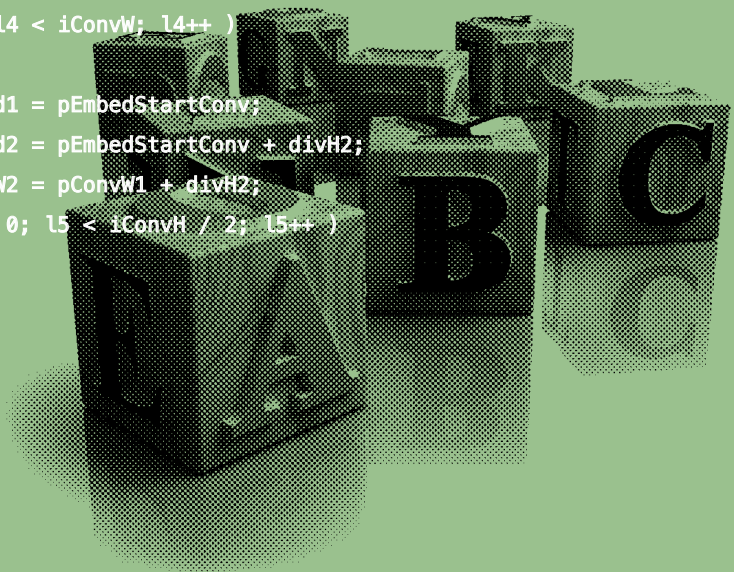
2018.04

AI CODE

```
// Calculate feature map for each out channels.
double* pConvWOutPointer = arrConvW;
double* pBias = arrConvB;

int divH2 = iConvH / 2;
double tmpreg = 0.0;
for( int l1 = 0; l1 < iConvOutChannel; l1++ )
{
    // 1. Calculate feature map for each output channel.
    double* pFeatureMap = arrFeatureMap;
    double* pEmbedStart = arrEmbedVec;
    *pPooled = -100.0;
    for( int l2 = 0; l2 < iConvResultW; l2++ )
    {
        double* pEmbedStartRow = pEmbedStart;
        for( int l3 = 0; l3 < iConvResultH; l3++ )
        {
            if( iConvH % 2 == 0 )
            { // Speedup-version
                double* pConvW1 = pConvWOutPointer;
                double* pEmbedStartConv = pEmbedStartRow;
                *pFeatureMap = 0;
                tmpreg = 0;
                for( int l4 = 0; l4 < iConvW; l4++ )
                {
                    double* pEmbed1 = pEmbedStartConv;
                    double* pEmbed2 = pEmbedStartConv + divH2;
                    double* pConvW2 = pConvW1 + divH2;
                    for( int l5 = 0; l5 < iConvH / 2; l5++ )
```

...



카카오에서 매일 발행하는 리포트입니다.

KAKAO AI REPORT

Vol. 12

발행일 | 2018년 04월 30일

발행처 | (주)카카오

발행인 | 여민수, 조수용

편집인 | 김대원

편집 | 김수정, 양현서, 윤위훈, 이수경,

인치원, 정하늬, 최은필

디자인 | 허진아

메일 | kakaoaireport@kakaocorp.com

브런치 주소 | <https://brunch.co.kr/magazine/kakaoaireport>

본 카카오 시리포트의 판권은 주식회사 카카오가 소유하고 있으며, 카카오는 카카오 시리포트를 상업적 목적으로 이용하는 것을 금지합니다. 또한 저작권법에 따른 정당한 범위 내에서 카카오 시리포트를 이용·가공·인용을 하시는 경우라도 반드시 출처를 밝혀주시기 바랍니다.

COVER

카카오 시리포트의 표지에선 AI와 관련된 의미 있는 코드들을 매월 소개하고 있습니다.

Vol.12 코드 | 최동현 heuristic.085@kakaocorp.com

카카오미니의 문장 분류기에서 사용하는

사용하는 CNN(convolutional neural network) 연산 알고리즘 일부를 발췌했습니다.

카카오미니의 문장 분류 알고리즘은

신경망(neural network)을 이용하여 기존의

알고리즘에 비해 큰 성능 향상을 이루었을 뿐만

아니라, 실 서비스 구현을 위하여 각종 최적화

기술이 적용되어 가볍고 빠르게 실행 가능합니다.

contents

preface 02

AI & life

일상 생활에 스며든 AI part.1

장은미 | 여가: AI가 쉼의 공간과 놀이의 시간에 스며들다 06

홍정우 | 노동: AI 시대에 직업의 의미 12

강영욱 | 소비: AI로 인한 구매와 유통 구조의 변화 18

Kakao inside

카카오미니 그리고 수학자와 AI

최동현 | 카카오미니의 명령어 분류 방법 **AI in Kakao** 24

임성빈 | 수학자는 어떻게 인공지능을 발전시키는가? 28

이수경 | 딥페이크와 리벤지 포르노 32

information

2018년 5월 AI 콘퍼런스 소개 38

closing

40

카카오 시리포트 12호를 내며

쌀쌀한 날씨를 비집고 피어났던 벚꽃 잎이 떨어지고 벌써 4월의 끝자락에 섰습니다. 벚꽃이 지면 새로운 꽃들이 찾아 오듯이 우리의 일상도 다양한 공간과 시간 속에서 변화를 맞이하고 있습니다.

특히 인공지능(artificial intelligence, AI)으로 인해 우리의 일상이 변화하고 있습니다. 어떤 분야는 잔걸음, 또 다른 분야는 큰 보폭으로 AI가 진입하고 있습니다. 편집진은 일상을 다섯 가지 형태(여가, 노동, 소비, 교육, 부동산)로 구분한 뒤, 각 분야의 전문가에게 AI가 변화시키고 있는 현재 양상과 앞으로의 변화 전망에 대한 글을 요청했습니다. 이 중, 이번 호에는 여가, 노동, 소비에 관한 글을 담았습니다.

여가를 위한 시간과 공간은 기술의 발전과 더불어 많은 변화를 거쳐 왔습니다. 최근의 여가를 위한 공간과 시간 변화의 기저에는 인공지능이 공통적으로 깔려 있습니다. 또 AI의 발전과 맞물려 주된 사회 문제로 간주되는 일자리 문제. 이번 호에 담기는 글에는 AI 시대에 노동의 의미에 대한 심도 깊은 분석이 담겨 있습니다. 그리고 현재 구매와 유통 단계에서 AI가 활발하게 적용되고 있는 양상은 ‘AI와 소비’란 키워드로 고찰된 글 안에 담았습니다.

카카오 인사이트 파트에는 카카오톡이 명령어 문장을 분류하는 방법에 이어 인공지능과 수학자의 불가분의 관계에 대한 글을 넣었습니다. 딥러닝 기술 발전의 그늘인 딥페이크(deepfake)와 페이크앱(fakeapp)을 다룬 글을 보시면서, AI 윤리에 대해 함께 고민하실 수 있는 시간이 되었으면 합니다.

이번 호에 유익한 글들을 기고해 주신 분들께 다시 한번 감사 인사를 드립니다. 편집진은 이번 호에 담긴 콘텐츠들이 많은 분들께 도움이 되기를 희망합니다.

2018년 04월 30일
카카오 시리포트 편집진 드림

일상 생활에 스며든 AI



part.1

AI & life

장은미 | 여가: AI가 삶의 공간과 놀이의 시간에 스며들다

06

홍정우 | 노동: AI 시대에 직업의 의미

12

강영욱 | 소비: AI로 인한 구매와 유통 구조의 변화

18

우리의 일상과 관련된 특별 테마들을 준비하였습니다. 먼저 이번호에서는 여가, 일, 소비와 관련된 내용을 우선적으로 담았습니다. 특히 여가 활동에 이미 깊숙하게 스며든 AI 기술들을 살펴보고, 앞으로 본격적인 AI 시대를 맞이하면서 눈 앞에 마주하게 될 직업시장의 변화와 우리의 대처 방안들을 살펴봅니다. 그리고 AI가 가져온 소비 생활의 변화에 대해서도 다양한 기업들의 적용 사례를 통해 알아보도록 하겠습니다.

여가

AI가 삶의 공간과 놀이의 시간에 스며들다

여가 공간은 한마디로 노는 곳이자 쉬는 곳이다. 어렵게 말하면, 여가 공간은 여가 활동이 이루어지는 장소 및 물리적 자원의 총체로서 생활의 구속에서 벗어나 편안하고 즐겁게 여가 선용 및 휴식을 취할 수 있는 모든 공간을 말한다. 최근에는 ‘한번 뿐인 인생’이라는 뜻의 ‘올로(you only live once, YOLO)’를 외치는 사람이 늘어나면서 여가 공간의 중요성 또한 점점 커져가고 있다.

여가 공간이란?

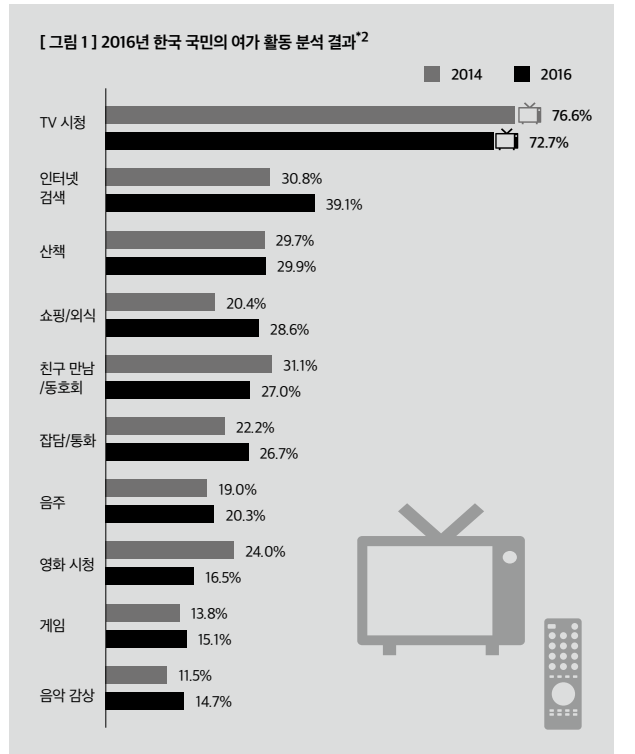
여가 활동은 개인의 경제적 능력, 가용한 여가 공간의 제공 여부, 노동시간 외에 여가 시간의 길이와 연속성에 따라 천양지차로 달라진다. 계절 별 골프 여행부터 ‘방에 콧 박힌다’는 의미의 ‘방콕’ 여행까지 개인의 경제적 수준과 성향이 다름에도 불구하고, 여러 신문과 방송 매체에서 ‘다가올 미래는 이럴 것이다, 혹은 저럴 것이다’라는 추측성 기사가 급증하였다. 대부분 중산층 젊은 남녀를 가정하여 제시된 시나리오를 읽어보면 ‘그럴 수 있겠구나’라며 긍정적인 공감을 보이기도 하지만, 말도 안 된다며 고개를 젓거나 개인의 취향이 다름에도 불구하고 강요받는 기분을 느끼기도 한다. 한편으로는 본인이 뒤처지는 것이 아닌가 하는 부정적인 모습까지 보이기도 한다. 그렇기 때문에 여가 공간에 미치는 인공지능의 영향에 대한 논의도 주관적일 수밖에 없다. 그럼에도 불구하고 여가 공간 전체를 다룰 수 있는 객관적인 논의가 가능할까?

사회가 변하면 노는 ‘것’도 달라지고 노는 ‘곳’도 달라진다

약 15년 전 대한민국에서 주 5일 근무제를 도입했을 때 발간되었던 수도권 주변의 여가 공간 실태와 개발 필요성을 연구한 보고서를 살펴보았다. 국내외 여가 활동 조사에서는 여가 활동 종류를 (1)관광/여행, (2)문예/문화, (3)체육/스포츠, (4)오락/위락, (5)취미/교양 중 행락과 등산과 산책, 낚시와 정원 가꾸기 (6)역사 탐방, (7)휴양, (8)사교, (9)TV시청 및 수면 등의 휴식, (10)자격증 공부 및 가족 모임 등 기타 활동으로 구분하였다.¹⁾ 주 5일 근무제가 실시됨에 따라 서울시민과 경기도민이 쉴 수 있는 여가 공간을 적극 만들어야 한다는 것이 보고서의 골자였다. 15년이 지난 지금의 시각으로 그때의 보고서를 다시 보면 여가 자체를 바라보는 눈이 바뀌었다는 것을 실감할 수 있다. (6)역사 탐방과 (8)사교는 대부분류에서 빠질 가능성이 높고 단순 인터넷 서핑이 대부분류로 들어갈 확률이 높다. 또한 (10)에서 언급된 기타 활동은 여가 활동에 포함되지 않을 수도 있다.

‘여가 시간’을 고려한 인공지능의 힘

2016년 국민여가활동조사 분석 결과 여가 시간을 보내는 순서를 정리하면 TV시청, 인터넷 검색, 산책, 쇼핑 및 외식, 친구 만남과 동호회 모임, 잡담과 통화, 영화보기, 게임, 음악 감상, 독서 순으로 나타났다.²⁾ 즉 2016년 말에는 TV리모콘이 한국인의 여가 시간을 쥐고 있음을 알 수 있다.



그렇다면 인공지능이 앞의 순위별 여가 활동에 어떻게 영향을 미칠까? 인공지능이 여가 활동에 영향을 미치려면 TV방송 채널의 선택에 영향을 주어야 하는데, 아직까지는 인공지능에게 채널 결정권까지 넘기지는 못한 상황이다. 하지만 일부 사람들은 리모콘으로 직접 조작하지 않고 “내가 좋아하는 영화채널 틀어줘”라는 음성 명령으로 해당 서비스를 이용하는 수준까지는 왔다. 그런 의미에서 카카오미니 스피커 및 통신사의 AI 스피커가 핫한 아이템이 되는 것은 당연하다. 그렇다면 인터넷 검색 및 게임 부분에 인공지능이 미치는 영향은 어떻게 될까?

검색 엔진에 인공지능을 적용해 온 지 벌써 10년이 넘었지만, 검색창 주변을 차지하는 광고와 관련 뉴스로 볼때 내가 무엇을 궁금해 하는지 남들이 다 알고 있다는 불편한 마음이 든다. 이러한 ‘불편함과 즉시적 만족성의 균형을 어떻게 맞출 것인가?’라는 고민에 슬그머니 답을 알려주는 과정을 넛지(nudge)라고 해야 할까?

여가 활동에서 9위를 차지할 정도로 게임은 우리 사회에 커다란 변화를 가져왔다. 게임의 시나리오와 실질적인 구현은 이제껏 사람에 의해 창조되어 왔다. 하지만 지금은 단순 그래픽 부분에 함수를 넣어 이미지와 동작을 생성시키는 표현 분야를 넘어서, 게임 자체를 시가 수행하도록 하는 실험, 그리고 시가 스스로 게임을 만들어 낼 수 있게 하는 저작 도구까지 고안하는 환경이 조성되고 있다. 전쟁과 전략 시뮬레이션 위주의 게임에서, 실제 도시를 계획하는 심시티(SimCity)를 넘어서, 민주주의(democracy) 혹은 전체주의(despotism)를 선택해 작은

글 | 장은미 emchang21@gmail.com

공간 정보와 IT기반의 컨설팅 회사를 운영하며, 대학에서는 비즈니스 GIS와 방재 GIS를 강의 하고 있다. GIS는 정보를 담아내는 그릇이라는 생각으로, 다양한 재해 관련 의사 결정 지원 시스템 구축 및 교육 훈련 체계를 구축해 왔으며, 산림에서 국토 정보를 넘어 수심측량과 해양지명에 이르기까지 기술 관련 어려운 이야기를 쉽게 풀어내기 위한 글쓰기를 즐겨한다. 현재 (주)지인컨설팅 대표이사, 서울시립대학교 겸임교수를 역임하고 있다.

마을을 관리하고 그 마을의 시장원리와 복지정책까지 구현하는 게임까지 만들어지고 있는 상황을 볼 때, 게임이란 중독되어서는 안 되는 것이란 관점에서 변화하고 있다.³ 게임 산업에서 인공지능을 고민하지 않을 수 없는 상태에 도래한 계기가 있다. 적지 않은 게임 고수들이 생겨나면서 게임 제작사의 시나리오를 모두 파악해버린 것이다. 게임을 만들어 내는 속도보다 스테이지를 달성하는 속도가 빨라졌고, 결국 자동 조합 모듈로 다양하고 복잡한 게임에 대한 요구가 생겨난 것이다.

네트워크 게임에서 플레이어가 직접 조정하지 않더라도 액션을 가지는 캐릭터들은 강화학습에 의해서 움직이게 된다. 자연어처리랩과 시랩이 게임 회사의 핵심 부서가 되면서 AI 전공 인력이 게임회사로 몰려가는 현상은 우리나라에서도 관찰된다.

산책과 걷기는 과도한 집중력을 요하지 않고 편안하게 걷는 산책과 다이어트나 근육 운동을 위한 파워워킹을 말한다. 스마트폰에 설정한 하루 목표 7,000보를 달성하게 되면 스마트폰 앱으로부터 칭찬을 받고, 하루 목표 이상의 걸음 수 알림을 보면서 사람들은 뿌듯한 마음으로 걷기 운동을 마칠 수 있게 된다. 개인의 걸음 수 측정 데이터를 SNS로 공유하기 시작한 것은 벌써 2~3년 정도 되었으니, ‘이제와서 뭘 새삼스레’하고 생각하는 수준에 이르렀다. 이제는 이러한 정보를 외부 보험사와 공유하면서 개인의 데이터 제공시 보험료 인하에 대한 선택 버튼이 생기고, 나의 심장 박동이 비정형 데이터로 저장되는 모습까지 그려질 정도다. 나이 드신 분들의 경우 벌써부터 필수기능이라고 좋아하시는 모습도 보이니, AI 기술을 좋은 방향으로 사용할 수 있다면 여가를 넘어 원격 의료까지 가능할 것으로 생각된다.

음악 감상은 어떨까? 라디오 혹은 카페에서 “어디서 많이 들었던 노래인데”라고 말하면, 조용히 하라며 앱을 켜고서 1분 이내에 가수와 노래 이름, 발표 연도를 읊어대는 모습을 종종 볼 수 있다. 이 모습에 놀라움을 느끼면서도 한편으로는 클래식 종류는 찾지 못 할 것이라고 확인했던 적이 있다. 실제로 유사한 멜로디를 추출하는 기능은 15년 전에 이미 개발이 완료되었으나 모든 음악이 아닌 특정 음악 시장을 대상으로 제한했던 까닭에 데이터베이스화도 불완전했다고 볼 수 있다. 최근에는 음원을 저장하고 인덱싱하는 방법에 대한 연구가 계속되고 있다. 예측하건대, 인공지능이 음악 감상이라는 여가 활동에 영향을 주려면 국가와 지역별로 서로 다른 문화적 차이를 고려하여 서비스가 제공되어야 할 것이다. 유사한 측면에서 지도 서비스도 정보 소비가 가능한 지역을 중심으로 시작하여, 서울 전체에서 대도시 권역으로, 이후에 전국으로 확대되고 최종적으로는 국경을 넘은 서비스 플랫폼으로 확대된 것이 아닐까 생각한다. 참고로 미국 시장에서의 음악 관련 AI 어플리케이션은 [표 1]처럼 요약될 수 있다.³⁴

해외의 리포트에 따르면, 이제 스마트폰의 사진 기능으로 많은 사람이 고품질의 사진을 찍듯이 일반인들도 대위법 및 다양한 변주 기법을 배우지 않고도 인공지능 기반의 앱으로 작곡을 할 수 있는 환경이 마련될 것이라고 한다.⁵

[표 1] 음악 작곡을 도와주는 인공지능 기반 앱(app)	
앱 이름	주요 기능
팝건(Popgun)	알리스(Alice)라는 플랫폼에서 연주자의 성향을 예측해, 인공지능으로 연주자와 같이 연주하거나 더 나은 연주를 위해 코치하면서 나만의 음악교실을 만드는 기능 제공
엠퍼(Amper)	클라우드 기반의 작곡 플랫폼으로 음악도서관에서 샘플 데이터베이스를 활용하여 작곡자가 부분음을 내면 유사한 음을 찾아 주면서 번이를 일으키고, 여러 옵션을 제공하면 사용자가 최종 선택을 해서 작곡을 완성하는 기능을 제공
스포티파이(Spotify)	사용자가 좋아하는 음악을 딥러닝 방식으로 필터링하여 유사한 성향의 그룹에 새로 나온 음악을 소개하고, 좋아할 만한 음악을 찾아내어 노출시켜 결과적으로 판매로 이어지게 함
판도라(Pandora)	150만 개 이상의 곡을 데이터베이스(database, DB)화하여 마신 리스닝 체계를 만들어 유사곡을 추천하며 개인 맞춤형 노래를 추천함. 지역별 다운로드 혹은 스트리밍 음악을 히트맵(hitmap)을 만들어 제공함으로써 가수들이 팬이 밀집된 곳에서 로드쇼를 할 수 있게 도와줌
랜더(Lander)	가수들이 적정가격에 음반 녹음을 할 수 있는 기능을 제공하여 신곡 발표 비용을 절감할 수 있음
클라우드바운스(CloudBounce)	음반 녹음부터 음악 소개까지 자동화된 풀 패키지를 제공하는 곳으로 랜더(Lander)보다 비쌌
성균관대학교	빅데이터를 세그멘테이션(segmentation)하여 작곡이 가능한 소프트웨어 기능을 제공

여섯 번째 순위의 잡담 및 통화는 카카오톡, 밴드 등의 SNS 활동을 포함한다. 아마 2017년에는 더욱 그 비중이 늘어났을 가능성이 있는데, 이는 장년층의 유입과 더불어 다양해진 SNS 채널과도 관련이 깊다. 매년 조사하는 카테고리 항목이 변경되기도 하므로 시간에 따른 추이 변화를 보기는 어렵지만, 잡담과 통화의 양식이 변경되었거나 같은 SNS 활동이라도 눈으로 읽기만 하는 활동도 잡담으로 볼 수 있는가에 대한 논의 역시 해보직하다.

앞서 말한 독서 활동의 경우, 상위 10가지 활동 가운데 마지막을 차지한다. 독서는 대형서점 및 인터넷을 통해 종이책, 전자책, 전자책 회원권 또는 오디오책 구매 등이 가능해지면서 그 형태가 다원화되고 있다. 책을 읽는 사람과 읽지 않는 사람들의 양극화는 점차 심화되는 가운데 도서 시장은 점점 작아지는 파이를 나누어야 하는 상황에 놓이면서 경쟁이 더욱 심해지고 있다. 하지만 시를 토대로 분석한 판매 성향이 소비자의 구매로 이어지도록 하고, 다양한 북 콘서트 및 큐레이션(curation) 서비스를 하면서 도서 시장이 점차 다변화 되어가는 것을 볼 수 있다. 또한, 시로 추천 받은 책, 그림 및 디자인은 카드뉴스를 통해 소비로까지 이어지게끔 만든다. 초기에 특허 정보와 같은 특정 서비스에서 필터링을 위해

개발되었던 기능이 이제는 책을 요약하고 키워드를 뽑아내는 등의 일상적인 도서 검색 및 내용 변환 서비스로 확대되어 가고 있다.

'여가 공간'을 고려한 인공지능의 힘

동일한 보고서의 여가 시간을 보내는 공간에 대한 조사 결과를 살펴보면 식당(9.4%), 아파트 내 집 주변 공터(8.7%), 생활권 공원(6.9%), 커피숍(6.1%), 대형마트(4.5%), 영화관(4.4%), 산(4.1%), 헬스클럽(3.8%), 목욕탕(3.3%), 기타 실내 공간(1.3%) 순서로 나타났다.

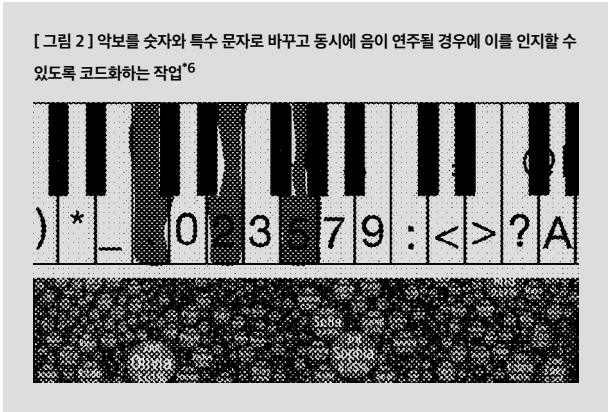
식당이 1위인 이유로는 '식사하셨어요?'라는 표현이 '안녕하세요'와 같은 인사말로 사용되는 우리나라의 문화와 맞물려 있다는 점, 그리고 최소 자본으로 전 국민의 집중도를 끌어냈던 식당 순회 프로그램과 요리 기반의 예능 프로그램 영향 때문으로 생각된다. 보는 것에 만족하지 않고 더 많은 정보를 찾아보고, 기록하고, 공유하며, 지인들과 식사를 하는 가운데 보내는 여가 시간의 중요성이 커지고 있고, 이는 단순히 다른 제품을 소비하는 것과는 차원이 다른 문제다.

두 번째 순위로 집과 아파트 등의 주변 공간 등을 꼽고 있는데, 이는 서울을 비롯한 대도시에서 집과 학교, 집과 일터를 중심으로 반복되는 생활 속에서 쌓인 피곤 때문에 집에 머무르면서 여가를 즐기려는 성향과 비슷한 맥락에 있다. 집은 개인에게 사적인 공간으로, 방은 잠을 자는 공간이기도 하지만 휴식과 같은 다양한 기능을 하는 공간으로 자리한다. 여기서 인공지능은 앞에서 확인된 TV 및 비디오 시청, 게임 등 개별화된 활동에 기여하고 데이터의 업로드와 다운로드 등은 수시로 전자제품과 서버, 그리고 휴대하는 여러 디바이스로 데이터 이동이 끊임없이 일어난다. 시각과 청각을 동시에 상호작용할 수 있는 비서 로봇의 역할이 커질 수 있다.

인공지능이 집에 영향을 주는 환경과 기존의 홈 오토메이션(home automation), 원격제어 등의 기존 IT 기술 사이에는 어떤 차이가 있을까? 몇 가지 질문으로 그 차이를 살펴해보도록 하자. 여가 활동과 관련된 영역에서 나의 키워드는 어디에 저장이 되는가? 그리고 분석하는 알고리즘은 누가 만드는가? 기존의 홈 오토메이션은 각자 개인이 결정한 것을 원격에서 조정하는 것과 세팅해 놓은 대로 조정되는 것인데 비해, 인공지능의 작동은 이와 달리 관련된 디바이스는 통신망과 서버를 통해 우리의 지능이 아닌 별도의 플랫폼 서버에 저장된다. 서비스를 받는 즉시 모든 사진과 검색 키워드, 구매 목록 등은 모든 것들이 센서에 연결되어 개인의 성향을 판별하게 된다.

무조건 “나는 시가 싫다”고 말하는 사람은 정보 노출을 최소화하는 삶을 선택할 수 있다. 아마존(Amazon)을 시작으로

사용자가 자주 사용하는 키워드와 사진을 분석하여 유사한 도서를 찾아주는 맞춤형 큐레이션이 등장하였고, 우리나라 인터넷 서점에서도 멤버로 가입한 회원에게 다운로드 받은 책들 혹은 독자가 좋아할 만한 도서 등을 알려주기도 한다. 또, 포털 메인 메뉴까지 맞춤형으로 구성하면서 개인의 여가 시간과 공간을 결정하는 정보를 최소의 비용으로 얻을 수 있게 해준다.



여행 관련 책자는 다른 분야에 비해 종이책으로 꾸준히 나오고 있으며 독립출판사의 발행 비율이 높다. 인공지능에게 무엇을 하고 놀 것인가를 물어보는 것은 종교를 가진 사람에게 “어떤 목사님이나 어떤 스님의 말씀을 들어볼까요?”라고 물어보는 것과 같게 느껴질 수 있다. 즉 인간으로서 주어진 짧은 여가 시간마저 선택을 기계에 맡기고 본인의 선택을 포기하는 모습을 거부하며, 비록 간접적으로 영향은 받지만 결정권은 각자에게 있다고 믿고 싶을 것이다.

하지만 이미 해외 여행 예약 사이트에서는 실제로 챗봇(chatbot)과 인공지능을 결합한 서비스를 선보이고 있다. 이렇게 하면 고객의 정보를 전자 메일, 페이스북(Facebook), 문자 메시지, 스케줄 등에서 취득하여 가본 곳과 가고 싶다고 말했던 곳, 또는 검색 기록을 기반으로 고객 맞춤형 여행 계획을 제안할 수 있다. 혹은 공항에서 갑자기 게이트가 바뀌거나 이변이 발생할 경우 자동으로 상황을 대처하는 방법을 알려줄 것이며, 여행이 끝난 후 여행 소감을 받아 다음 여행자에게 어떻게 하면 더 나은 서비스를 제공할 수 있는지에 대한 데이터를 모으고, 정보를 제공해 준 여행자에게 리워드 포인트로 선물을 남겨주기도 할 것이다.

네덜란드 항공(KLM)에서는 이미 챗봇 시스템을 도입하였으며, 음악이 시와 접목됐을 때와 유사하게 여행 상품 혹은 같이 갈 동료나 친구 추천, 관련 정보와 비디오의 제공, 성향과 기분 상태 모니터링, 고객 관리 등에 활용하고 있다. 이후에는 더 효율적인 디지털 광고 대상 여부를 판단하거나 챗봇의 기능을 개선하는 데 사용할 것이라 보고 있다.⁷

글을 마치며

정보통신기술 이전에 예술과 학문이 어떠한 형태로든 먼저 존재했음은 분명하다. 하지만 소통의 주요 도구가 정보통신기술이 되면서, 웹을 통해 공유하지 않으면 존재하지 않는 것과 마찬가지로는 논쟁이 불기 시작했다. 또 다시 실존주의 논쟁이 살아난 것 같다. 문학 연구에 정보통신기술이 적용된 사례를 추정한 연구에서도 미래의 예술과 학문에 대한 변화에 IT 기술이 독이 될 수도 있고, 약이 될 수 있음을 보여주고 있다.*8

‘나는 옛 것을 고집하고, 내 여가시간을 보낼 때만큼은 기계에 의존하지 않고 살리라’하고 마음을 먹은 사람이라면 최근의 변화로 인한 영향을 덜 받을 수 있다. 하지만 우리의 주변인들이 스마트폰을 사용해 여러가지 편리한 서비스를 즐기기 시작하면서 우리의 여가 공간에 조금씩 시가 스며들게 되었다. 예전에 한 이동 통신사의 광고에서 나온 말이 생각난다. “또 다른 세상을 만날 때는 잠시 꺼두셔도 좋습니다”고 말하지만, 다시 네트워크에 접속하는 순간 우리의 여가 공간은 여러 명의 빅브라더(big brother)*9와 함께 하는 공간이 되어버릴 것만 같다.

충청북도 괴산의 산막이 길을 지날 때, 비목(樵木) 팻말을 본 사람 중 여럿은 ‘초연히 쓸고 간 깊은 계곡 양지 녘에’로 시작하는 가곡을 부른다. 하지만 갑자기 휴대폰에서 혹은 주변의 전자 기기에서 그 노래가 나온다면 무섭지 않을까? 또 다른 케이스로는 ‘삼미슈퍼스타 마지막 팬클럽’ 전자책 소설을 보고 있는데 ‘연안부두 떠나가는 배’ 노래가 휴대폰에서 휴대폰이 울리면 어떨까? 내 뇌를 스캔하는 빅브라더에 대한 의식때문에 휴식 공간에서만이라도 누군가의 간섭 없이 편히 쉬고 싶은 마음이 들 수도 있다.

마지막으로, 앞에서 잠깐 언급했듯이 아이돌 가수의 노래는 찾아도 클래식인 경우 데이터베이스가 방대하여 찾지 못 할 것이라고 했지만, 인공지능으로는 가능할 수도 있다. 바흐(Bach) 악보를 미디(Musical Instrument Digital Interface, MIDI) 파일로 저장한 후 CSV(comma separated value) 파일로 변환하고, 이것을 다시 알고리즘으로 분석하여 머신 리스닝 한다.*6 이 파일을 훈련 모드를 반복하여 악기로 연주하기도 하며, 훈련 시간을 늘리면서 그 악보를 기초로 멋진 재즈로 변화시키기도 한다. 훈련 시간이 너무 길어지면 오히려 이상한 음악으로 변하기도 하니 인공지능도 적정 수준의 인풋(input)을 가져야 한다는 원리가 맞아 들어가는 것 같다.

우리가 살아가는 동네, 뛰어 놀던 공터가 사라지고 놀이터는 비어가면서, 이제는 승용차에게 그 공간을 내어주고 있다. 우리에게 여가 공간이라는 것은 집이나 카페 같은 편한 곳으로 인식된다. 카페의 백색소음과 공부하는 사람들 사이에서, 같은 테이블에 마주보고 앉아 있어도 조용히 카카오톡으로 대화하는 그 곳이 지금의 여가 공간이 되고 있다. 하지만 앞으로는 자연의 소리를 듣고

체험하며 묵언의 공간에서 여가 시간을 보내는 등 지금까지와는 전혀 다른 방향의 여가 공간을 찾아 달려갈 수도 있다. 독자의 권리 중 하나가 떠오른다. 책을 읽지 않을 권리, 건너뛰며 읽을 권리, 끝까지 읽지 않을 권리, 다시 읽을 권리, 아무 책이나 읽을 권리, 아무 데서나 읽을 권리, 군데군데 골라 읽을 권리, 소리 내서 읽을 권리, 읽고 나서 아무 말도 하지 않을 권리가 그것이다.*10 이제는 여가 공간에 슬며시 찾아 든 시에게도 같은 권리를 주장할 수 있다. 이러한 주장을 하고 넋지를 놓는 일이 여가 분야에서 AI 산업의 핵심 목표가 되지 않을까?!”

*1 참고 | 윤양수, 김의식, 2002, 레저행태 변화와 여가공간 조성방안연구, 국토연구원 *2 참고 | 2016 국민여가 활동조사, 한국문화관광연구원 *3 참고 | 이경혁, 2017, 게임, 세상을 보는 또 하나의 창, 로고폴리스 *4 참고 | https://www.techemergence.com/musical-artificial-intelligence-6-applications-of-ai-for-audio/ *5 참고 | http://www.marketexpress.in/2017/10/will-ai-change-the-future-of-music.html *6 참고 | 머신 리스닝 설명 애니메이션 https://www.youtube.com/watch?v=5acogDL_4JU *7 참고 | https://www.webcredible.com/blog/artificial-intelligence-tourism-travel/ *8 참고 | 장은미 박용재, 2017, 정보통신기술의 적용이 문학연구에 약이 될 것인가? 독이 될 것인가? 한국문학연구 제 55집 *9 설명 | 정보의 독점으로 사회를 통제하는 관리 권력, 혹은 그러한 사회체계를 일컫는 말. 사회학적 통찰과 풍자로 유명한 영국의 소설가 조지 오웰(George Orwell, 1903~1950)의 소설 《1984년》에서 비롯된 용어이다. 긍정적 의미로는 선의 목적으로 사회를 돌보는 보호적 감시, 부정적 의미로는 음모론에 입각한 권력자들의 사회통제의 수단을 말한다. *10 참고 | 다니엘 페니아크, 2004, 소설처럼, 이정임 옮김, 문학과 지성사 *11 참고 | 리처드 세일러 캐스 선스터인, 2009, 넋지(nudge), 안진환 옮김, 리더스북

노동

AI 시대에 직업의 의미

이 글이 주로 돌아다닐 네트워크 세상인 통신이나 미디어 분야의 시각에서 인공지능이라는 주제를 다룬다면 애플(Apple)의 시리(Siri) 혹은 구글(Google)과 아마존(Amazon)이 내세우는 서비스, 카카오와 국내 통신사들이 판매에 매진하고 있는 대화형 인공지능과 자동차 등의 자율주행과 관련한 지능화, 상품 유통 부문과 엔터테인먼트, 서비스 영역의 디지털 지능화와 자동화가 주류가 될 것이다. 이 주제들은 비교적 일상과 관련한 일들인지라 잘 알려져 있고 상대적으로 이해하기도 쉽다. 이런 주제들을 중심으로 다루면 살짝 부족한 느낌이 들기 때문에 우리나라가 세계 경제에서 차지하고 있는 위치라든지, 그 자리에 오르기까지 중요한 역할을 했던 다른 제조업 분야들을 감안해서 이야기를 끌어가는 것이 전체적인 그림을 이해하는 데 도움이 될 것 같다. 그리고 미래 상황을 예지하는 내용은 기술로 인해 사람과 시장이 어떻게 연관지어 동작하는 지에 대해 살펴보고, 그런 상황들 속에서 어떤 방향으로 논의가 발전하게 될 것인가에 대해 추정해 보는 것이 유용하다. 직업의 방향도 그 영향을 받는 것이 일반적이기 때문이다.

기술 발전에 따른 변화를 바라보기 위한 시각

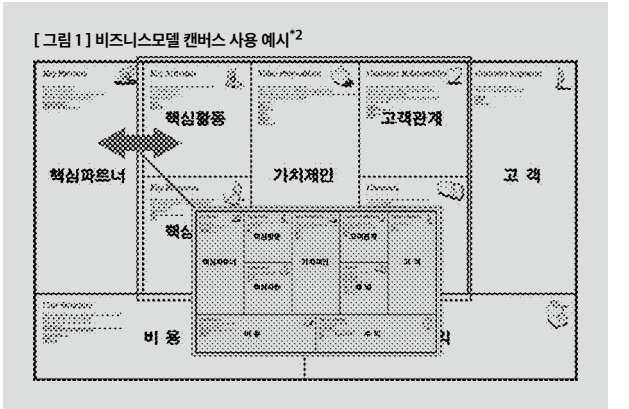
새로운 기술은 과거의 기술이 어떠한 방식으로든 노동과 결합하여 부가가치 창출 과정에 도약적인 효용성을 제공함으로써 기존에 비해 차별적인 부가가치를 생산하게 한다. 이것은 기존의 시장 경쟁 관계 균형에 영향을 주는 것으로 생각되어 왔는데, 실제로는 경쟁 관계뿐만 아니라 협력 관계 혹은 공생 관계에도 영향을 미치게 된다고 보는 것이 좀 더 타당하다. 이는 원천기술의 개발부터 최종 시장의 상품으로 소비되는 과정을 어떤 단계로 구분하는가에 따라 다른 시각으로 평가할 수 있고, 분업이 기본인 현대사회에서 단위 분업에서 이루어지는 기술혁신의 효과에 대해서 평가하고자 한다면 조금 더 미시적인 시각이 필요할 수 있기 때문이다.

주식회사라는 개념과 소액주주운동의 활성화를 통해 자본주이자 고용주가 되는 경우도 존재할 수 있다는 점도 이런 시각이 유효할 수 있다는 근거가 된다. 기술이 사회에 어떤 영향을 미치게 될 것인가에 대한 고민은 1차 산업혁명 시기를 살았던 이들이 기록으로 남겨 후세의 사람들이 여러 가지 방향으로 고민해 보도록 만들었다. 그 중에서 가장 큰 영향을 남겼다고 할 만한 마르크스의 논의가 주로 노동 가치에 대한 것이라는 인식이 일상적이지만, 그 ‘노동 가치론’ 사상은 당시의 기계와 분업, 공장 시스템에 대해서 선구적인 연구를 했던 앤드류 유어(Andrew Ure), 찰스 배비지(Charles Babbage), 아담 스미스(Adam Smith)의 영향을 많이 받았다고 알려져 있다.

자본주의 사회의 동인인 더 많은 이윤을 추구하는 행위가 기술 발전을 촉진하며, 그 과정에서 벌어진 자본주와 노동자들간의 투쟁에서 자본주는 노동자들의 영향력을 줄일 수 있는 방안으로 기술을 활용하는 방안을 떠올렸다. 하지만 마르크스는 자본주의가 만들어 내는 계급 구조에서의 기술은 자본주가 노동자를 억압하는 수단으로 활용될 수 있지만, 계급이 타파된 사회주의 사회에서의 기술은 노동자의 지겨운 육체노동을 해방시키는 수단이 될 수도 있다고 설명한다.¹⁾

어떠한 부가가치가 탄생하는 과정을 비즈니스라는 시각으로 바라본다면 비즈니스가 성립하기 위해 어떤 요소들이 존재하는지 볼 수 있고, 그 요소들 간의 연계 관계에 대한 이해를 통해 비즈니스가 어떠한 모습을 취하고 있는지 살펴 볼 수 있다. 시라는 자동화, 효율화 기술들이 발전하고 확대 적용됨에 따라 일자리가 어떻게 변화할 것인가에 대해 논의하기 위해서는 해당 기술이 적용되어 변화하는 비즈니스의 수요와 공급관계가 어떻게 변화하게 될 것인가를 가장 우선으로 생각해 봐야 한다. 이런 분석을 위한 도구로 [그림 1]의 비즈니스모델 캔버스²⁾ 방식을 사용하기도 한다. 비즈니스모델 캔버스를 사용하면 캔버스 요소들 중에 등장하는 행위 주제들 사이의 관계에 대해서도 동일한 캔버스 모델을

재귀적인 방식으로 적용하며 생각해 볼 수 있다. 어떠한 주체가 건전한 사업을 영위하며 사회적 활동 주체로써 존재하기 위해서는 비즈니스 모델이 합리적이어야 한다. 특히 비즈니스의 수익과 지출의 비용구조가 불합리할 경우, 모든 것을 소비하여야만 존재 가능한 세상 구조에서는 지속 가능성이 낮을 수 밖에 없다.



시를 도입한 기업들의 반응

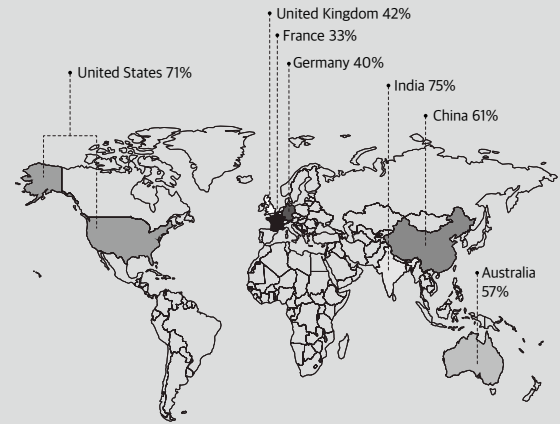
시는 주요 정보화 기술들 중에서 기존의 생산 기술에 데이터를 기반으로 인간의 고등 학습 기능과 행동 양식들을 부여하여 효율 극대화를 추구하는 기술이라고 정의 가능하다. 시는 생산 장비 자체의 구조적 설계 영역이라든지 물성개선 등의 영역에 적용할 수 있으나, 주된 적용은 기존 시장에서 가장 큰 효용성을 발휘할 수 있는 영역에서부터 시작하는 것이 상식적인 접근이다. 속된 표현을 쓴다면 많은 돈을 만들 수 있는 영역에의 적용이 우선시 된다고 볼 수 있다.

인도에서 시작하여 ‘가장 존경받는 지식기업상(most admired knowledge enterprises, MAKE)’을 수차례 수상해 온 글로벌 IT컨설팅 회사인 인포시스社は, 시를 대표하는 정보화 기술의 각 사회에 미치는 영향성에 대한 보고서를 지난 몇 년간 발간해 왔다. 한국과 일본을 제외한 미국, 영국, 중국, 독일, 프랑스, 호주, 인도의 거주자 20만 명을 대상으로 조사한 결과들이다. 여러 나라의 기업들을 대상으로 매년 이뤄지는 조사를 통해서 그 변화가 어떤 영향을 미치고 있으며 어떤 지향점을 가지고 있는지에 대해 살펴 볼 수 있다. 2018년 보고서는 ‘Leadership in the Age of AI : Adapting, Investing and Reskilling to Work Alongside AI’라는 타이틀로 발간되었다.³⁾

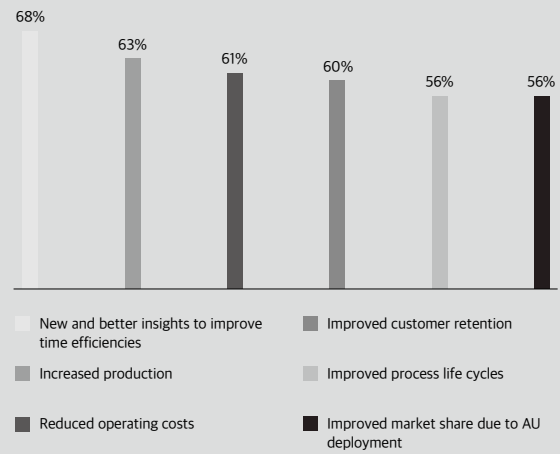
글 | 홍정우 jwhong@kisti.re.kr

2009년 영화 ‘국가대표’를 제작할 때 슈퍼컴퓨터를 사용하여 특수효과를 만들었다는 소식이 우리나라 인터넷 뉴스 포털을 채운 적이 있다. 당시에 잘 알려지지 않은 연구/개발 영역에서 쓰이던 프로페셔널 디지털 자원과 도구들을, IT클라우드 서비스 형태로 제공해 온 슈퍼컴퓨팅 환경에 적용해서 비전문가들이 첨단 디지털 연구와 개발 자원을 직접 설정하고 원격 활용하게 할 수 있다는 개념을 만들었다. 그리고 영화 ‘국가대표’의 제작과정에서 해당 개념의 일부를 시연했었고, 3D프린터와 같은 첨단기술 장비들과 인력들을 클라우드 서비스화 한다는 이야기가 매스컴을 오르내린 적도 있었다. 그런 연구를 하고 시연을 했던 까닭은 바로 디지털 기술 발전으로 인한 위해 혁신적 변화가 어떻게 사회 경제적으로 다가올 것인가에 대한 고찰과 대응 방안에 대한 고민의 발로였다. 그 고민들을 바탕으로 첨단 디지털기술을 제품디자인 분야에 적용해서 이제껏 존재하지 않았던 개념들을 응용한 상품을 개발하는 일을 하다가 지금은 정부출연연구원이 주도하는 원천기술 사업화라는 개념을 정립하기 위한 실험을 여전히 하고 있다. 이 시대는 혁신역량 경쟁 시대라고 할 수 있기도 하지만, 그 혁신의 결과가 결국 첨단기술을 이용해서 새로운 시장과 고용 기회를 창출하는 것으로 이어지는 것이 대응책들 중 한가지라고 생각했기 때문이다. 실무 구현하여 증명하는 것을 우선하는 현장 엔지니어링을 견지하고 싶기 때문에 정확한 이론을 근거로 하고 있지는 않다는 것을 전제로 하고 싶다. 미래 시장은 남보다 먼저 상상해내고 선정하며 만들어내야 하는 것이라고 생각하기 때문이다. 어디서 읽었는지 기억은 나지 않지만 추격자가 따라오지 못할 만큼 더 빨리 세상을 바꾸자는 것이 오바마 정부의 과학기술과 산업정책이었다. 현재 KISTI 한국과학기술정보원 슈퍼컴센터 소속으로 3D프린팅산업협회이사를 역임하고 있으며 에너지, 환경 문제 등에 관심을 가지고 있다.

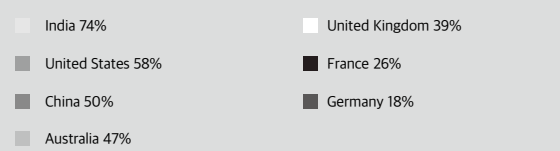
[그림 2] 성공적으로 AI 투자 효과를 이끌어 내고 있는 국가들³



[그림 3] AI를 기반으로 한 디지털화가 기업에게 미친 영향³



[그림 4] 국가별 AI가 특정 업무에서 사람을 능가한다고 생각하는 비율³



세계 각국이 어떻게 대응하고 있는지 통계 데이터를 통해 살펴 볼 때는 각국의 역사와 사회 정치적 분위기와 현재 산업과 경제 구조 등을 감안하는 것이 필수적이다. 디지털화라는 기술적인 변화가 어떤 국가와 사회에 자리잡고 있는 모습을 살펴 보는 것은 해당 변화로 인해 영향을 받는 수많은 이해관계들과 문제점들이 어떻게 조율되고 해결되었을까 하는 물음에 대한 답을 찾는 행위라고 할 수 있다.

어떤 경영자라도 실제 기술 투자를 통해 수익 개선 등의 효용성이 발생 가능한 영역에 투자를 하는 것이 상식이다. 이는 최소한 기술 관련한 부분을 배제하고도 시장 경쟁력이 있거나 기술

개선이 있을 경우 시장 경쟁력을 가지는 사업 모델을 가진다는 의미이기도 하다. 인포시스 리포트에서 발췌한 [그림 4]에서 각국의 시장, 인건비와 노동 규제 및 소비 상품 등을 고려하며 살펴보면 몇 가지가 추정 가능하다. 미국은 자국 내 제조업을 되살려 내려고 하고 있고, 인도는 수많은 인력 자원을 활용한 산업들이 많다고 알려져 있다. 또 중국의 경우 국가가 공산주의 체제를 통해서 인민들을 보호하고, 기업들에게 이들의 보호를 의무적으로 요구하는 일이 드물다. 하지만 중국식 공산주의 체제 하에서 빈부 격차가 심화되는 것은 사회 체제상 허용되기 어렵다.

각 국가의 정책자들은 국내 시장을 지켜내면서 해당 국가의 고용을 지속하면서 자국 시장을 유지하는 정책을 펼치게 된다. 기업은 자발적으로 혹은 노조와의 협의를 통해, 혹은 정부가 제시하는 정책에 따라서 기존 고용 인력들을 다른 업무에 배치하는 등의 대응을 하게 된다. 농업과 같은 분야에서는 어떤 상황이 벌어질까? 국내가 소농 위주의 방식인 것과는 다르게 넓은 토지를 가진 국가들은 대부분 대규모 자동화 농업 방식으로 운영된다. 물론 자본이 농업을 지배하고 있다고 알려져 있다. 해당 국가에서 농업에 종사하는 사람들은 자본에 고용된 엔지니어로 분류되어질 가능성이 높다는 뜻이다. 디지털화되고 자동화된 농업은 자본 의존적이게 됐다.

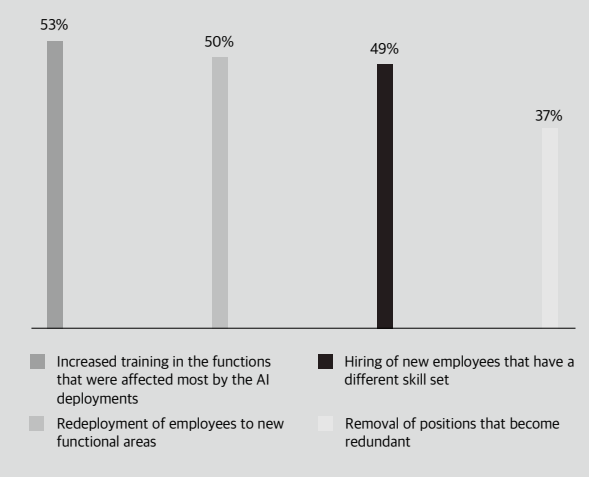
대규모 디지털 자동화 농경을 추진하고 있는 글로벌 농산자본과의 가격 경쟁에서 상대적 약위에 놓일 가능성이 높은 각국의 소농들은 소득 저하 문제를 겪을 가능성이 높다. 이 문제는 전통적인 도농 소득 격차로 인한 농민들의 도시 노동자화 문제로 이어질 가능성이 높을 뿐만 아니라, 각국이 디지털 자동화 농경을 추진한다든지 하는 방식으로 대응하지 못할 경우 식량 수급 안보라는 이슈로 발전하게 되며 지역에 따른 고용 기회 창출 불균형 문제를 악화시키는 큰 요인이 된다.

공산주의 혹은 사회주의 성향을 띠는 국가에서는 국가가 자본을 만들어 농업을 전략적으로 통제할 수도 있다. 낮은 농업생산량을 보이고 있는 국가에는 농민들에게 자동화 농업기술을 어떻게 저렴하게 보급할 것인가 하는 것이 더 큰 과제로 자리잡게 된다. 이것은 농업 부문에서 예상되는 업태 변화다.

[그림 5] 일본 구보타사의 완전자동화 농기계 개발 전략 발표 보도자료⁴(왼쪽)와 17명의 제봉사 뭉을 하는 AI 기반 자동 봉제기계 소개 보도자료⁵(오른쪽)



[그림 6] AI 도입에 따른 각 기업의 인력 부문 변화 및 대응 현황³



[그림 6]은 AI 도입으로 사업 내에서 어떤 형태의 조정을 했나 라는 질문에 대한 답이다. [그림 2]에서 각 나라를 표현한 색상을 보면, 미국과 호주는 재교육과 재배치를, 인도는 일자리를 줄인다고 대답한 것을 알 수 있다. 프랑스는 AI 도입으로 신규 일자리가 창출되고 있는 효과를 보여준다. 신규 고용은 새로운 경쟁력 확보를 위한 것이라면, 재교육과 재배치는 자의 혹은 타의로 직원들을 보호하는 행위로 분류할 수 있다. 법과 실제 제도가 적용되는 모습들은 해당 사회가 지향하고자 하는 것을 함의하고 있다고 봐도 무리가 없다. 이는 직업들이 변화하며 국가가 지향하는 바를 파악하는데 도움이 된다. 세계 분업구조 속에서의 역할과 분배 형태를 함께 생각하는 것도 미래 상황을 예측하는데 유용하다. 중간재 생산과 수출 비중이 높고, 세계 교역 순위 10위 안에 위치하고 있는 우리나라의 경우 어떤 인력 정책을 펼치고 있는지에 대한 질문도 상당히 의미미하다.

AI를 도입하고자 하는 분야들과 그 효과

디지털 혁신기술들 중 주목 받고 있는 디지털트윈 기술은 유지보수 현장 인력이 데이터가 지시하는 대로만 작업을 하면 되므로 현장에 투입되는 고급 인력은 필요 없어질 가능성이 높다. 이 기술을 사용하면 유지 보수를 빙자한 기술 유출 방지도 가능하고, 인건비도 대폭 절감 할 수 있기 때문에, 기업에게는 다양한 장점을 제공하는 기술이라고 볼 수 있다.

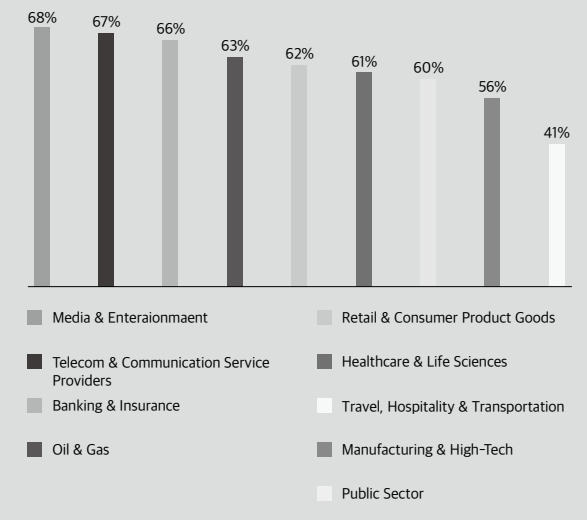
하지만 피고용 기회를 찾는 입장에서 보상 수준이 좋은 일자리를 얻기 위한 진입 장벽의 높이와 그 위치를 유지하기 위해 필요한 신기술 획득 노력이 어떤 수준일 것인가 하는 이슈, 그리고 진입을 가능하게 하는 훈련 프로그램의 제공 주체와 이와 관련한 획득 용이성이 문제가 발생한다.

[그림 7] 디지털트윈 개념도⁶



AI는 주요 정보화기술들 중에서도 기존 생산 기술에 데이터를 기반으로 한 인간의 고등 학습기능과 행동 양식들을 부여하여 효율 극대화를 추구하는 소프트웨어화 기술이라고 정의내리는 것도 가능하다. AI는 방대한 영역에 적용 가능하며 컴퓨터 응용 영역에서 다양한 연구들이 진행되어 왔다.

[그림 8] AI 적용을 주도할 인력 확보에 어려움을 겪고 있다고 대답한 분야³



[그림 8]에서 AI 도입에 어려움을 겪고 있는 분야들은 미디어와 엔터테인먼트, 통신과 통신 서비스 분야, 은행과 보험, 오일 & 가스, 소매와 일반소비재, 건강관리와 라이프 사이언스(life science), 여행/접객/운송, 제조 & 첨단기술, 공공분야 등 다양하다. 다양한 분야에 AI가 도입되었을 때 누리게 될 효과들은 [그림 3]에서 논의되었던 것들이 아닐까 생각된다. 실제로 해당 분야에서는 새로운 아이템 발굴, 효율성, 생산성, 비용 축소, 고객 유지, 공정

개선, 시장 확대 등의 효과가 있었다고 이야기한다. 이는 동일한 생산량이라는 전제 조건에서 현재와 비교하여 시장 경쟁력 강화와 고용 축소로 인한 인력 효율화 효과가 발생했다는 의미다. 생산을 늘이기 위해서는 시장 확대가 선행되어야 하지만, 고용 축소로 소비 감소가 예상되는 상황에서 시장 소비가 늘어나는 것을 기대하기란 쉽지 않다.

여러 가지 대응 방안들

인포시스의 리포트는 AI 기술이 직접 와닿지 않는 수사적인 그 무엇에서 벗어나서 어딘가에 효과를 발휘하고 있다고 이야기하고 있다. 이러한 상황을 벗어나기 위해서 첫 번째로 인간 우선의 경영, 두 번째로 평생 공부하는 분위기를 만드는 것, 세 번째로 사업의 모든 영역에 대해 투명성을 기할 것, 마지막으로 비즈니스 프로세스의 자동화 그 너머의 가치를 고민할 것을 제안했다. 그렇다면 그 이유에는 무엇이 있을까 라는 질문을 던지고 우리의 상황과 비교하며 답을 찾는 것이 필요하다.

각국의 경쟁 기업들은 AI 기술의 확산과 적용 영역의 확대로 경쟁의 무게 중심이 첨단 인력들이 창출하는 혁신 역량에 의해 좌우된다는 점을 깨닫고 있다. 실질적인 체험을 중요하게 취급하는 교육의 방향성과 오픈 소스화, 온라인 접근성 강화 등의 경향은 전 세계로부터 인력을 흡수하기 위한 기반으로 동작한다. 또한 기업들은 첨단 인력을 먼저 확보하기 위해 앞다투어 나서고 있다. 글로벌 산업 구조 속에서 우리의 역할은 제조기지라고 자평해왔다. 그 효율성이 보장될 수 있었던 것은 인력 양성 체계부터 평가, 분배 전체가 유기적으로 동작해왔기 때문이라는 평가가 있다. 그러나 산업화 시대의 효율성은 창의성과 인간을 기반으로 하는 혁신적 효율성과는 괴리가 큰 개념이다.

당연히 우리나라 내부에서 창출되는 고용 기회와 해외에서 만들어지는 기회는 시간 차이를 두고 선행되고 후행하는 모습을 보일 것이지만, 이는 결국 언젠가는 동기화 될 가능성이 높다. 디지털화는 노동과 자본의 균형에서 자본 쪽에 더 유리하게 만드는 효과를 만들어 내면서 각국은 시장을 유지하기 위해서 첨단기술의 확산 방안을 강구할 것이다. 그리고 기술로 인해 창출되는 부가가치의 분배에 대해 더 고민할 수 밖에 없다. 물론 그 과정에서 생산 잠재력이 있는 국토와 자원을 가진 나라가 유리하다는 것은 당연한 사실이다.

시로 인한 미래 직업 변화에 대한 논의는 궁극적으로 자본과 노동, 기술혁신으로 창출된 가치의 분배 이슈들을 다룰 수 밖에 없다. 그렇게 해야 혁신으로 인해 소멸한 과거의 것들을 대체할 수 있는 새로운 생산과 소비가 창출될 기회가 만들어지고, 새로운 시대를 맞아 확대 재생산될 수 있기 때문이다.

“인간과 기계가 짝을 이룸으로써

인간이 지금 잘하는 일을 ‘더’ 잘할 수 있게 해준다면,

마찬가지로 기계가 지금 하는 일을 훨씬 더 잘할 수 있게 해준다면?

그렇다면 그것은 ‘증강’이 될 것이다.

단순히 노동의 분할을 넘어 가치의 증식이 될 것이다.”

<AI 시대 인간과 일> 中
(토머스 데이븐포트, 줄리아 커비 저)

* 이 글에서 참고 인용한 통계 자료들은 인포시스의 2018년 보고서 'Leadership in the Age of AI : Adapting, Investing and Reskilling to Work Alongside AI'로 부터 가져온 것이다.

*1 참고 | 홍성욱, (2006) [기술속 사상] "기회는 괴물이다", 한겨레 <http://www.hani.co.kr/arti/culture/book/153335.html#csidxcad0ef94155932581870c9948da0eef> *2 참고 | 비즈니스모델 캔버스 <https://brunch.co.kr/@givemore/3> *3 참고 | Leadership in the Age of AI : Adapting, Investing and Reskilling to Work Alongside AI (<https://www.infosys.com/age-of-ai/Documents/age-of-ai-infosys-research-report.pdf>) *4 참고 | http://www.newsis.com/view?id=NISX20180226_0000237133&clD=10101&pID=10100 *5 참고 | <https://www.geek.com/tech/this-robot-can-sew-t-shirts-as-fast-as-a-crew-of-17-humans-1714383/> *6 참고 | 인더스트리 4.0과 디지털트윈 (제조업이 천생연분을 만나다) (https://www2.deloitte.com/content/dam/Deloitte/kr/Documents/insights/deloitte-anjin-review/09/kr_insights_deloitte-anjin-review-09_08.pdf)

소비

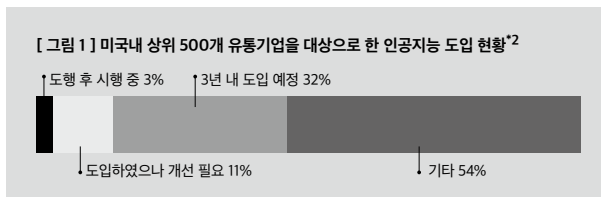
AI로 인한 구매와 유통 구조의 변화

최근 정보통신기술의 급격한 발전은 소비자들의 구매 방식에도 변화를 야기시키고 있다. 소비자들은 과거와 같이 여러 유통 단계를 거친 전통적인 구매 방식이 아닌, 해외 판매자로부터 직접구매(직구)를 하거나 공동구매 플랫폼을 통해 도매업자 혹은 제조업자로부터 직접구매(공구)를 하고 있다. 물건을 구매하는 채널도 오프라인에서 온라인과 모바일로 급격히 확장되고 있다. 특히 스마트폰의 등장으로 시공간 제약이 줄어들면서 모바일 쇼핑이 온라인 쇼핑을 견인함으로 인해 오프라인 유통 업체 역시 온라인 채널을 확대하고 있다. 온라인 쇼핑의 확대와 오프라인 업체의 온라인 시장 진입으로 두 가지 쇼핑 방식의 경계가 모호해지면서 온·오프라인을 넘나드는 크로스쇼퍼(cross shopper)가 증가하고 있다.

유통업의 변화

온라인과 모바일을 통한 상품의 다양화와 상품에 대한 과잉 정보 제공은 소비자에게는 구매 결정을 어렵게 하는 상황을 만들고 있다. 이에 따라 유통 업체에서는 인공지능, 빅데이터 등 정보통신기술을 활용함으로써 소비자가 필요로 하는 상품을 쉽게 검색하고, 소비자의 취향을 분석한 후 소비자 성향에 맞는 상품을 추천한다. 그리고 필요로 하는 상품을 적재적소에 제공하는 방식의 서비스를 확대하고 있다. 미국 유통업 전문 조사 기관인 보스턴 리테일 파트너스(Boston Retail Partners, BRP)가 2017년 미국 유통기업을 대상으로 조사한 결과에 따르면 40% 이상의 기업들이 챗봇(chatbot)이나 인공지능 비서 등 인공지능 기술을 도입했거나 3년 내에 도입할 계획이라고 한다.¹⁾

챗봇(chatbot)
채팅(chatting)과 로봇(robot)의 합성어로 사람과 대화를 통해 질문에 알맞은 답이나 각종 연관 정보를 제공하는 인공지능 기반의 커뮤니케이션 소프트웨어를 지칭한다.



유통업에서 인공지능 접목이 가능한 분야

급격하게 변화하는 유통 분야에서 현재와는 다른 검색 엔진이나 쇼핑 로봇이 필요하다는 것은 여러 연구에서 설명되고 있는데, 이 가운데 인공지능을 접목하는 것이 가장 강력하고 유망한 분야라고 언급하고 있다.³⁾⁴⁾ 유통업의 변화에서도 기술된 바와 같이 최근 젊은 층을 중심으로 쇼핑 접근성, 상품의 다양성, 가격 비교, 전자 결제의 편리성 등으로 온라인 유통 채널이 각광받고 있다. 소비자와의 접점이라 할 수 있는 검색 단계에서 인공지능 기술이 접목된다면 어떠한 변화가 가능할까? 일반 검색 엔진의 예를 들어보자. 고객의 검색어에서 핵심 키워드를 분석해 정보를 검색하고 그 결과를 고객에게 제공하지만, 인공지능 기술은 고객의 특성, 선호도, 특별한 요구 등 고객 상황을 스스로 인지한 후 해당 정보에 기반하여 상품을 선택하고 추천해줄 수 있을 것이다. 또한 유통사업자는 재고 여부와 고객의 현재 위치 등을 인식하여 실시간으로 고객 맞춤형 서비스를 제공할 수 있을 것이다.

유통 분야에 인공지능 기술을 어떻게 접목할 수 있을지는 세계 각국의 인공지능 기술 개발 관련 어플리케이션을 살펴보면 어느 정도 가능할 수 있다. 가장 대표적인 예로

IBM의 ‘왓슨(Watson)’, 구글 딥마인드(Google-DeepMind)의 ‘알파고(Alpha-Go)’, 페이스북(Facebook)의 AI 기반 인공지능 로봇 ‘챗봇(Chatbot)’, 일본 소프트뱅크(SoftBank)의 가사도우미 ‘페퍼(Pepper)’ 등을 들 수 있다. IBM은 인공지능 컴퓨터 ‘왓슨’이 쇼핑 어드바이저, 스마트 머천다이징, 판매 전략 수립, 유통 업체 내부 업무 효율화, 와인 어드바이저, 건강 어드바이저, 레시피 어드바이저 등의 서비스 영역에 적용될 수 있을 것으로 예상하고 있는데, 이 가운데 쇼핑 어드바이저 서비스가 가장 유망할 것으로 분석하고 있다.⁵⁾

인공지능 적용사례

1) 노스페이스의 왓슨 적용

아웃도어 브랜드인 노스페이스(The North Face)는 2015년 12월에 IBM의 인공지능 슈퍼 컴퓨터인 ‘왓슨’을 기반으로, 커머스 솔루션 업체 플루이드(Fluid)의 소프트웨어 ‘엑스퍼트 퍼스널 쇼퍼(expert personal shopper, XPS)’를 기반으로 하는 양방향 개인화 온라인 쇼핑 플랫폼 ‘The North Face XPS’를 출시하였다. 이 플랫폼은 왓슨이 제공하는 인공지능 기술을 활용하여 온라인 쇼핑몰의 이용객이 마치 서비스 담당자와 대화하듯이 짧은 시간에 원하는 상품을 찾도록 하는 서비스를 제공한다. 이 플랫폼에서 왓슨은 상호 작용을 통한 지속적 학습을 통해 시간이 경과할수록 고유의 가치관과 지식을 습득하는 특징을 가지고 있다. 왓슨은 자연어 처리를 통해 문법 및 맥락을 파악하고, 가능한 모든 의미를 평가하여 복잡한 질문의 의미를 파악한 뒤 방대한 자료에 근간하여 최적의 상품을 제시한다. ‘XPS’는 최초의 인공지능 커머스 플랫폼으로 평가받고 있다. 현재 왓슨은 자연스러운 대화를 이끄는 점에서 한계를 표출하고는 있지만, 이 같은 미비점을 개선할 경우 개별 브랜드의 특성에 따라 상이한 성격을 가진 인공지능 퍼스널 쇼퍼로 기능하며 온라인 쇼핑 경험을 혁신할 수 있을 것으로 전망된다.

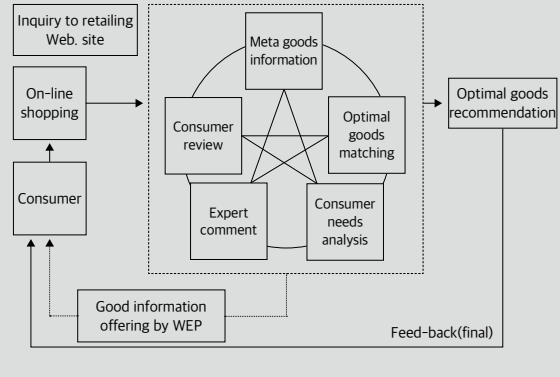
IBM 인공지능 ‘왓슨’의 쇼핑 정보 지원 서비스
1997년 ‘딥 블루(Deep Blue)’라는 슈퍼 컴퓨터를 만든 IBM은 2011년 ‘왓슨’이라는 이름의 인공지능 컴퓨터를 미국 TV퀴즈쇼 ‘제퍼디(Jeopardy)’에 출연 시켜 인공지능 기술을 세계에 널리 알리는 계기를 만든다. 이후 2014년 1월 아웃도어 전문기업 노스페이스와 함께 뉴욕에서 개최된 미국소매협회(National Retail Federation, NRF) 행사에서 인공지능 기술을 적용한 XPS(expert personal shopper) WEP 프로토타입인 쇼핑 정보 지원 서비스 사례를 발표한다.
왓슨의 쇼핑 정보 지원 서비스는 [그림 2]와 같다. 백팩을 구매하고자 하는 고객이 있다고 가정하자. ①고객이 먼저 구매하고자 하는 제품을 IBM 왓슨 컴퓨터에 문의하면 ② XPS WEP은 소비자에게 필요한 다양한 정보(등산지, 체류기간, 현지 날씨 등)를 분석하고 ③ 소비자에게 최적의 제품을 추천함과 동시에 부가적으로 다양한 제품 정보들을 보여 준다. ④ XPS WEP은 고객이 당초 구매하고자 했던 배낭과 함께 침낭, 모자 등 최적의 제품 정보들을 함께 추천·제공함으로써 고객이 최적의 상품을 추천받을 수 있을 뿐 아니라 부가적으로 필요한 관련 상품 정보도 제공받아 구매까지 이루어지도록 도와준다.

글 | 강영욱 ykang@ewha.ac.kr

공간 정보 빅데이터 분석 및 지오비주얼라이제이션(geovisualization)에 관심이 많다. 최근 공간정보 데이터 분석에 시간 차원을 포함한 데이터 분석 및 시각화, 데이터마이닝, 텍스트마이닝 기술을 접목하면서 자연스럽게 인공지능 핵심 요소 기술의 변화와 접목에도 관심을 갖게 되었다. 유통이나 소비 분야 전문가는 아니지만 AI 기술이 격전지를 이루는 소비 분야의 최근 동향을 살펴보고, 자료를 정리하는 차원에서 원고 청탁을 수락하게 되었다.

IBM은 2014년 4월 유통산업 관련 '퍼스널 쇼핑 도우미(personal shopping assistant)' 시스템을 구축하기 위해 온라인 전자상거래 소프트웨어 업체인 '플루이드'에 투자하였고, '왓슨'의 상용화를 위해 금융과 의료 분야에 적극 나서고 있다. 한국 IBM은 2016년 롯데백화점과 공동으로 인공지능 쇼핑 도우미 개발에 나서기로 하였다.

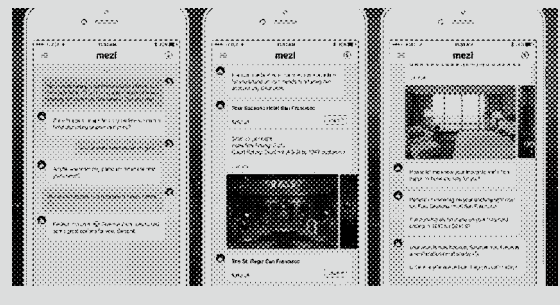
[그림 2] IBM 왓슨을 이용한 노스페이스의 백팩 구매 과정⁶



2) 인공지능 퍼스널 여행 어시스턴트 앱 '메지'

2015년 12월 스타트업인 메지(Mezi)는 문자를 주고받는 방식의 인터페이스로 온라인 쇼핑을 지원하는 인공지능 퍼스널 쇼핑 어시스턴트 앱 '메지'를 출시하였고, 2016년 후반부터 여행에 초점을 두고 있다. 메지 앱은 메시지 스타일의 인터페이스를 통해 사용자가 여행하고자 하는 목적지의 항공, 숙박, 레스토랑을 예약할 수 있는데, 이 과정에서 자체적으로 보유하고 있는 딥러닝과 자연어처리 기술을 적용하고 있다. 여행 앱과 관련해서는 카약(Kayak), 구글(Google)과 같은 대기업부터 히트리스트(Hitlist), 스카이스캐너(Skyscanner)와 같은 스타트업까지 다양한 앱들이 있다. 하지만 메지는 사용자가 개인정보 제공과 지불 방법을 결정하면 거래까지 한번에 이루어지는 간편함을 제공한다는 측면에서 차별성을 가지고 있다.⁷

[그림 3] 메지 앱으로 여행 관련 예약을 하는 모습⁷



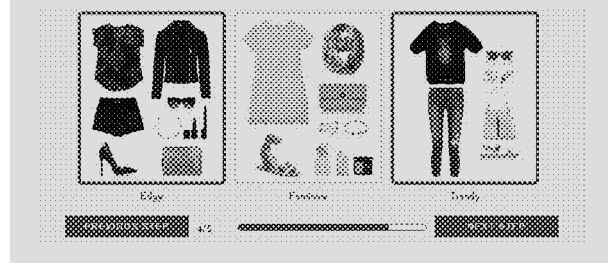
3) 여성 의류 온라인 퍼스널 쇼핑 서비스 '스티치 픽스'

2011년에 설립된 여성 의류 온라인 퍼스널 쇼핑 서비스 '스티치 픽스(Stitch Fix)'는 이용자가 자신의 기호에 맞는 스타일 프로필을 작성하면 독자적인 빅데이터 분석 알고리즘을 통해

전문 스타일리스트의 선정을 거친 컬렉션을 이용자 거주지로 배송하는 서비스를 제공한다. '스티치 픽스'는 2014년 6월 넷플릭스(Netflix)의 빅데이터 분석 담당자였던 제프 매그너슨(Jeff Magnusson)을 데이터 플랫폼 책임자로 영입하였다.

'스티치 픽스'는 총 5개의 의류 및 액세서리로 구성된 상자를 원하는 날짜에 맞춰 무료로 배송하며, 개별 상품 가격은 통상 55달러 수준이지만 이용자들이 별도로 자신이 구매하고자 하는 상품 가격 범위를 설정할 수 있도록 한다. 상품 선정을 위한 스타일링 비용으로 20달러를 청구하며 각 의류에 관련된 스타일 활용법을 포함하는 것이다. 해당 서비스는 3일 동안 구매 여부를 확정된 후 실제 구매한 상품에 대해서만 결제하면 되고 5개 상품을 모두 구매할 경우 25%의 할인이 적용된다. 마음에 들지 않는 의류는 선불 봉투에 싸서 반송하는 방식이다.⁸

[그림 4] 스티치 픽스 웹사이트에서의 상품 추천 서비스 예시⁹



4) 인공지능 가상 개인 비서 서비스

가상 개인 비서(virtual personal assistants) 서비스는 '머신러닝, 음성인식, 텍스트 분석, 상황 인지 등 인공지능 기술과 첨단기술의 결합으로 사용자의 언어를 이해하고 사용자가 원하는 지시 사항을 수행하는 소프트웨어 어플리케이션'으로 정의할 수 있다. 스마트폰, 웨어러블 디바이스 등 각종 IT기기 등과 결합하여 자연스럽게 사용자의 말을 이해하고 사용자가 원하는 정보, 작업을 알아서 수행하는 것이 가능하다.

현재 가상 개인 비서 서비스 시장에는 애플(Apple)의 시리(Siri), 구글의 나우(Now), 마이크로소프트(Microsoft)의 코타나(Cortana), 아마존(Amazon)의 에코(Echo) 등 기존 정보통신기술 기업들의 서비스와 에이미(Amy), 비브(Viv) 등 스타트업들의 신규 서비스가 존재하고 있다. 인공지능 비서 기능은 저마다 특색이 있는데, 애플 시리의 경우 사용자의 말을 이해하는 능력이 뛰어나고, 마이크로소프트의 코타나는 교통 정보 업데이트나 일정표 관리 등 기본적인 커뮤니케이션과 정보 수집 기능이 탁월하다. 한편 구글 나우는 막강한 검색 기능과 노래 제목을 알아내는 능력이 뛰어나다. 아마존의 에코 스피커에 탑재되어 있는 알렉사(Alexa)의 경우 가정용 서비스에 특화되어 있다.⁸

5) 이미지 기반 검색 서비스

과거 소비자들은 온라인에서 검색을 위해 텍스트를 입력했지만, 최신 스마트폰의 막강한 카메라 기능은 사용자 하여금 '카메라를 키보드'로 사용하는 것을 가능하게 한다. 인공지능에 기반한 이미지 검색은 아직은 초기 단계이지만 온라인에서 소비자가 본 이미지나 소셜 네트워크 서비스(social networking service, SNS)에서 '좋아요(like)'로 표시한 이미지에 머신러닝 기법을 적용하여 소비자가 무엇을 선호하는지 이해하고 이에 근거하여 추천을 하는 서비스를 제공한다. 소비자는 본인이 원하는 것을 타이핑하는 것이 아니라 원하는 아이템의 사진을 검색 엔진에 올리고, 여기에 가격 정보가 함께 있다면 가격대에 따라 검색하거나 본인이 원하는 가격대만으로 검색할 수도 있다.¹⁰

미국의 백화점인 노드스트롬(Nordstrom)이 운영하고 있는 '트렁크 클럽(Trunk Club)'이라는 온라인 의류 서비스 사이트는 인공지능 기반의 이미지 검색 서비스 기능을 강화하기 위해 핀터레스트(Pinterest)에 투자를 확대하였다.

핀터레스트(Pinterest)

핀터레스트는 이미지 공유 및 검색 사이트다. 이 서비스의 가장 큰 특징은 마음에 드는 이미지 파일을 '핀 잇(pin it)'하여 자신이 만든 보드에 모으고 관리할 수 있다는 점이다. 다수의 핀터레스트 사용자들은 최신 유행을 좇기 위해 서비스를 이용하는 것으로 알려져 있다. 핀터레스트는 큐레이션(curation) 서비스를 제공해 사용자가 자신이 사고 싶은 제품의 이미지를 모으는 데서 그치지 않고, 이미지를 클릭하여 해당 사이트로 이동해 제품을 구매하게 만든다. 이러한 특징 때문에 기업에서는 핀터레스트를 주요 SNS 마케팅 채널과 새로운 유통채널로 활용하기도 한다. 특히 이미지 콘텐츠가 주류인 패션 브랜드에서 핀터레스트를 마케팅 수단으로 사용한 사례가 많다. 2014년의 경우 가장 팔로워(follower)가 많은 패션 관련 브랜드는 엘엘빈(L.L.Bean)과 노드스트롬(Nordstrom)이었다.¹¹

[그림 5] 웨스트엘름: 이미지 기반으로 검색하고 구매하도록 지원하는 가구 브랜드¹²



미국의 가구 브랜드인 웨스트엘름(West Elm)은 사용자들이 오프라인 매장을 방문하지 않고도 SNS에 올린 피드백과 이미지를 보고 구매할 수 있게 하였다. 핀터레스트와 협업하여 웨스트 엘름의 가구를 구매한 소비자가 올린 내용을 확인하고 구매까지

가능하도록 기능을 구현한 것이다.

인공지능에 기반한 이미지 검색은 온라인 쇼핑 영역에서 크게 성장할 분야로 기대되고 있다. 유통업자들은 온라인 소비 시장이 거의 이미지에 기반하여 검색을 하므로 인공지능이 온라인 마케팅을 실행하는데 있어 핵심요소가 될 것으로 예견하고 있다. 한편 유통 업체 측면에서는 기존에 대규모 유통 업체들이 시장을 주도하고 있기는 하지만, 신기술을 빠르게 적용한다면 유통 분야에서 스타트업이 성장할 수 있는 기회가 될 것으로 분석하고 있다.

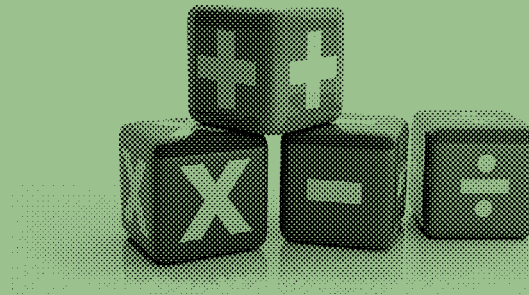
소비 분야의 인공지능 접목이 가져올 미래

인공지능은 글로벌 소비자들의 경제 생활에 큰 변화를 가져올 것으로 예상된다. 똑똑한 인공지능 쇼핑 지원 도구들이 속속 등장하면서 쇼핑이 한층 더 즐거운 경험으로 변모할 것으로 보인다. 원하는 가격과 스펙은 물론 최신 트렌드와 내밀한 감성·욕망까지 포착한 후 최적의 상품을 제안하는 지능적 쇼핑 도구들을 통해, 소비자들에게 인공지능 소비의 새로운 경험을 안겨주는 시대가 도래할 것으로 예상된다. 이미 아마존 등 많은 전자상거래 업체들이 인공지능의 기계학습 기법에 기반한 사용자 맞춤형 추천 기능을 운영하고 있지만, 인공지능 기법의 진화와 더불어 이런 추천 제안 기능은 더욱 정교하고 지능적인 모습으로 바뀔 것이라 전망된다.

¹ 참고 | 전해영, 2017, 4차 산업혁명에 따른 유통업의 변화, VIP Report, 현대경제연구원, 통권 710호. ² 참고 | Boston Retail Partners. (2017). 2017 Customer Experience Unified Commerce Survey. ³ 참고 | Chung, J. B. (2017). Internet Shopping Optimization Problem With Delivery Constraints. Journal of Distribution Science, 15(2),15-20. doi:10.15722/jds.15.2.201702.15 ⁴ 참고 | Stamford, C. (2015). Gartner's 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations That Organizations Should Monitor. Gartner Newsroom, Retrieved from http://www.gartner.com/newsroom/id/3114217 ⁵ 참고 | IBM IBV Report (2015). 2015 Consumer Purchasing Behavior Analysis Report at South Korea-China-Japan. Retrieved from IBM https://www-03.ibm.com/press/krko/press/release/4761-0.wss#release ⁶ 논문 | 김혜경·김원기, 2017, 인공지능 쇼핑정보 서비스에 관한 탐색적 연구, Journal of Distribution Science 15(4): 69-80 ⁷ 참고 | Lora Kolodny (2016). AI-powered virtual assistant, Mezi, pivots to focus on travel, https://techcrunch.com/2016/11/15/ai-powered-virtual-assistant-mezi-pivots-to-focus-on-travel/ ⁸ 참고 | IRS Global, 2016, 4차 산업혁명을 주도하는 인공지능 기술의 시장동향 및 주요 이슈 종합분석. ⁹ 참고 | https://www.stitchfix.com/ ¹⁰ 참고 | Cortney Harding (2017). How AI Is Transforming the Shopping Experience Based on the Images Consumers Look at Online, http://www.adweek.com/digital/how-ai-is-transforming-the-shopping-experience-based-on-the-images-consumers-look-at-online/ ¹¹ 출처 | 위키피디아. https://ko.wikipedia.org/wiki/핀터레스트 (2018년 3월 4일 검색) ¹² 참고 | http://mobilemarketingmagazine.com/facebook-collection-retail-ad-format-lifestyle-templates-catalogue

카카오미니

그리고



수학자와 AI

Kakao inside

최동현 | 카카오미니의 명령어 분류 방법 AI in Kakao

24

임성빈 | 수학자는 어떻게 인공지능을 발전시키는가?

28

이수경 | 딥페이크와 리벤지 포르노

32

카카오 인사이드에서는 카카오의 내부 이야기들을 담았습니다. 카카오미니가 음성 명령어 문장을 인식하고 분류하는 방법을 소개합니다. AI 기술 발전에 많은 기여를 해온 수학 이론과 수학자들을 알아보고 AI와 수학이라는 학문의 불가분 관계를 살펴봅니다. 그리고 음성과 영상 합성 기술을 통해 현재 우리 사회에 만연한 거짓과 조작의 이슈 및 이러한 사회적 문제들의 해결방안 모색에 대한 내용을 준비하였습니다.

카카오미니의 명령어 분류 방법

사용자가 카카오미니에게 명령을 내렸을 때 카카오미니로 입력된 음성 신호는 먼저 텍스트 형태로 변환된다(음성인식). 변환된 텍스트는 어떤 분야와 관련된 기능을 수행하기 위한 명령인지 식별된 뒤(도메인(domain)/봇(bot) 분류), 해당 분야의 기능 중에서 어떤 특정 기능을 수행하기 위한 것인지 판별된다(인텐트(intent) 분류).

문장 분류와 카카오미니의 관계

카카오미니의 음성 신호의 변환 단계(ex. “SBS 라디오 틀어 줘”)

(1) 음성인식

- 텍스트 형태인 ‘SBS 라디오 틀어 줘’로 변환

(2) 봇 분류

- 라디오와 관련된 기능 요청임을 인지

(3) 인텐트 분류

- 라디오를 틀어 달라는 요청임을 파악(라디오를 멈추거나, 가능한 라디오 채널 목록을 알려 주거나, 지금 듣고 있는 라디오 채널의 이름을 들려 주는 게 아니다)

(4) 슬롯 추출

- 정확히 어떤 라디오 채널을 요청하는 것인지 파악(이 경우에는 ‘SBS’)

(5) 요청 동작 수행

- 실제로 해당 라디오 채널의 주파수 및 재생과 관련된 정보를 찾아와 카카오미니 클라이언트에 전달

이 글에서는 (2)~(3) 단계에 대해 중점적으로 다룰 예정이다.

(2)와 (3) 단계는 주어진 문장을 여러 개의 분류 가능한 ‘클래스(class)’ 중 하나로 분류하는 것을 공통적인 문제로 정의할 수 있다. 예를 들어, (2) 봇 분류의 경우 어떤 특정 문장이 주어졌을 때 여러 경우의 봇 중에서 하나의 봇을 선택하는 문제로 볼 수 있다. 카카오미니의 경우 음악/라디오/뉴스/주식/인물/실시간 검색어 등의 봇이 있다. (3) 인텐트 분류의 경우 어떤 특정 문장이 주어졌을 때, 라디오 재생/라디오 정지/라디오 채널 목록/현재 듣는 라디오 채널명 반환 등 라디오 봇과 관련된 여러 기능 중 하나를 선택하는 문제로 볼 수 있다.

두 단계가 동일한 문제로 정의될 수 있다면, 왜 두 단계를 하나로 합치지 않는지 궁금해 할 수 있다. 그 이유는 크게 두 가지다. 첫째, 서로 다른 봇에서는 동일한 명령어도 서로 다른 동작이 이루어지도록 해야 하기 때문이다. ‘재시작’이라는 명령어는 음악 봇에서는 현재 듣고 있는 음악을 반복 재생 해주어야 하지만, 알람 봇에서는 이전 타이머를 재시작 해주어야 한다. 둘째, 특정 명령은 단독으로 사용될 경우 해당 봇에서 실행되면 오류이지만, 해당 봇의 맥락(context)에서는 정상적으로 동작해야 하기 때문이다. 예를 들어, “정답은 코끼리야”라고 사용자가 단독으로 말했을 경우 카카오미니는 어떠한 기능도 수행하면 안된다. 하지만 사용자가 카카오미니와 스무고개 중이었다면 미니는 이 말을 정상적인 명령으로 받아들여 동작한다.

이러한 종류의 문장 분류 문제는 언뜻 보기에는 굉장히 쉽게 해결될 수 있는 것처럼 보인다. 대부분의 사람들은 특정 문장을 보았을 때 그 의미를 직관적으로 이해할 수 있기 때문이다. 그래서 보통 키워드나 패턴 기반의 문장 분류 문제를 풀고자 한다. 그러나 이러한 방법들은 아래와 같은 두 가지 이유로 대부분 실패한다.

첫 번째, 어떤 문장에 결정적인 키워드가 포함되어 있더라도 주변의 맥락에 따라 문장 전체의 의미는 얼마든지 달라질 수 있다.

이는 아래 발화의 예제에서 잘 드러난다.

키워드 기반의 문장 분류 예시

“뉴스 라디오 틀어줘” → 라디오에서 뉴스 틀어주기(라디오 봇)

“라디오 뉴스 틀어줘” → 라디오에 관련된 뉴스를 틀어주기(뉴스 봇)

“라디오로 뉴스 틀어줘” → 라디오에서 뉴스 틀어주기(라디오 봇)

두 번째, 패턴 기반일 경우 서로 다른 봇과 관련된 패턴간 충돌이 다수 발생한다(ex. “윤종신 나이 들려줘” → 인물 VS 음악). 또한, 패턴 기반일 경우 발화의 변동에 매우 취약하다. 본 시스템의 입력은 사람이 직접 타이핑하는 것이 아닌, 음성인식 모듈을 거쳐서 들어오는 입력 값이기때 (어느 정도의 오류를 포함할 수 있음) 발화의 변동에 취약하다. 카카오미니에서는 문장 분류를 위하여, CNN(convolutional neural network)*을 이용하는 기계학습 기반 접근 방법을 취했다. CNN은 수많은 이미지 관련 문제를 해결하는데 성공적으로 적용된 인공지능망 구조로, 네트워크에 들어온 입력 벡터를 미리 선역된 수용 영역(receptive field) 범위만큼의 단위로 묶어서 처리한다. 이를 통해 입력에서 너무 세부적이어서 분류에 있어 중요하지 않은 부분은 제거하고 실제로 분류에 중요한 부분들만 도출할 수 있도록 설계되어 있다.

문장 분류를 CNN에 적용하기 위해서는, 먼저 사용자의 발화를 벡터 형태로 변환시켜 CNN 입력으로 사용 가능한 형태로 만들어야 한다.

발화를 벡터로 표현하기

발화를 벡터로 표현하기 위해서 본 작업에서는 먼저 발화를 이루는 각 단어들을 해당 단어를 나타내는 벡터와 맵핑(mapping)한다. 그 후, 주어진 발화를 구성 단어들을 나타내는 벡터로 변환한다.

단어를 벡터로 바꾸는 방법 중 잘 알려진 것들로는 원-핫 인코딩(one-hot encoding), Word2Vec, GloVe, FastText 등이 있다. 원-핫 인코딩 및 Word2Vec에 대해서는 카카오 시리포트의 이전 호²에 간략히 설명이 되어 있으므로 GloVe와 FastText에 대해서만 간략히 설명한다.

GloVe³는 워드 임베딩 벡터(word embedding vector)로 단어 i와 j의 차이를 표현하기 위하여, ‘프로브 워드(probe word)’ k를 사용한다. [표 1]은 i = ‘ice’, j = ‘steam’ 일 때, 프로브 워드 k를 사용하여 두 단어의 의미를 구분하는 예시다.

글 | 최동현 heuristic.085@kakaocorp.com

어느새 30대 중반으로 접어든, 그러나 마음만은 젊게 지내는 인공지능 개발자입니다. 작년에 카카오투 이직하며 카카오미니 개발에 참여하게 되었습니다. 카카오미니에 기여를 할 수 있어서 매우 기쁘고 미니를 좀 더 똑똑하게 만들기 위하여 열심히 공부 중에 있습니다.

[표 1] 대상 단어 'ice' 와 'steam'에 대한 프로브 워드 k의 동시 발생 확률 및 그 비율³

Probability and Ratio	k=solid	k=gas	k=water	k=fashion
P(k ice)	1.9 × 10 ⁻⁴	6.6 × 10 ⁻⁵	3.0 × 10 ⁻³	1.7 × 10 ⁻⁵
P(k steam)	2.2 × 10 ⁻⁵	7.8 × 10 ⁻⁴	2.2 × 10 ⁻³	1.8 × 10 ⁻⁵
P(k ice) / P(k steam)	8.9	8.5 × 10 ⁻²	1.36	0.96

즉, 단어 'ice'와 단어 'steam'의 차이를 표현하기 위해서 다양한 프로브 워드 k를 이용하여 동시 발생 확률(co-occurrence probability)의 비율(ratio)을 구한 후, 이를 이용해서 임베딩 벡터를 업데이트하는 과정을 거친다.

FastText⁴는 Word2Vec의 변형으로, 각 단어에 대한 임베딩 벡터를 도출하는 Word2Vec과 달리, 하나의 단어를 여러 n-gram의 그룹으로 보고, 각 n-gram⁵의 임베딩을 계산한다. 단어의 임베딩은 해당 단어가 포함하는 n-gram들의 임베딩의 조합으로 계산된다.

단어 '안녕하세요'를 예로 들면, n=3일 경우 해당 단어는 그룹 {'<안녕', '안녕하', '녕하세', '하세요', '세요>'}으로 정의되고, Word2Vec의 방식을 활용하여 각 n-gram에 대한 임베딩 벡터를 계산하게 된다. 각 단어의 임베딩 벡터는 해당 단어를 이루는 n-gram들의 임베딩 벡터의 조합으로 결정된다. 따라서 FastText 방식은 훈련 시 말뭉치에 존재하지 않았던 단어에 대해서도 알려진 n-gram으로 나누어 분석함으로써 임베딩 벡터를 도출할 수 있다. 이는 문법적 요소가 n-gram을 통해 분석되므로 한국어의 존댓말 표현과 같은 문법적 유사도를 파악하는 데 큰 강점이 있으나, 의미적 유사도를 파악하는 데에는 Word2Vec와 비교하여 성능이 약간 떨어지게 된다.

카카오미니에서의 발화 처리는 문법적인 요소보다 의미적인 요소가 훨씬 중요하게 작용하므로 미니의 문장 분류기에 사용된 단어 임베딩 벡터의 훈련에는 GloVe 알고리즘이 사용되었으며, 임베딩 벡터는 형태소 단위로 훈련되었다. 평가를 위해 9,736 쌍의 단어 유추(word analogy) 평가 데이터 셋이 자체 구축되었는데, 이러한 평가 셋은 {단어 A}:{단어 B} = {단어 C}:{단어 D}의 형태를 띄며, w(B) - w(A) + w(C) 가 w(D)와 가장 가까운 단어 벡터인지를 판별한다. 평가 데이터 셋의 예로는, 가평:경기 = 군산:전북(도시와, 도시가 속한 도의 관계), 밥:진지 = 딸:따님(높임말 관계) 등이 있다.

또한, 영어와는 다른 한국어 어순의 특징 때문에(영어: 주어-동사-목적어, 한국어: 주어-목적어-동사), 영어에서는 쉽게 판별되는 주어-동사간의 동시 발생 정보가 한국어에서는 주어-동사간의 거리가 인접하지 않으므로 판별력이 떨어진다. 이 문제를 해결하기 위해 카카오에서 자체 보유한 한국어 의존 문법 기반 구문 분석기(dependency parser)를 이용하였다. 기존

GloVe 알고리즘에서는 문장 내에서 단어와 단어 간의 거리를 측정하였다면, 의존 구문 문법 파서를 이용한 방법은 의존 구문 트리 내에서의 거리를 측정한다.

[표 2]는 다음 뉴스 2017년 1월부터 6월까지(총 14G)를 이용하여 훈련한 단어 벡터에 대한 단어 유추 셋의 평가 결과이다. TOP 20%의 빈도(frequency)를 가지는 단어에 대해서만 수치를 계산하였다.

[표 2] 구문 분석 트리를 이용한 GloVe 임베딩 벡터 도출 시 성능 향상

방식	# Correct	# Known	# Total	Known %	Total %
GloVe	3,763			46.10	38.65
GloVe + parsetree	3,816	8,162	9,736	46.75	39.19

위의 결과에서 보듯이, 일반적으로 낮은 구문 분석기의 분석 성능에도 불구하고 실제 단어 유추 말뭉치(word analogy corpus)에서 약간의 성능 향상이 있음을 알 수 있다. 따라서 위의 가설을 더욱 잘 반영할 수 있는 동시 발생 빈도(co-occurrence frequency)를 계산하는 것이 현재 해결해야 할 숙제다. [표 3]은 사용되는 데이터의 분량을 늘렸을 때의 실험 결과이다.

[표 3] 말뭉치 크기에 따른 GloVe 임베딩 벡터 성능 향상 추이

데이터	# Correct	# Known	# Total	Known %	Total %
14G(2017~)	3,816			46.75	39.19
41G(2016~)	4,426	8,162	9,736	54.23	45.46
69G(2015~)	4,492			55.04	46.14

문장 분류를 위한 신경망(neural network)

실제 네트워크의 입력으로는 위의 단어 벡터 외에 아래와 같은 특성들의 임베딩 벡터가 추가로 입력된다.

1) Local word vector

GloVe로 얻어진 단어의 의미벡터는 일반적인 상황에서의 단어의 의미를 나타낸다. 그러나 단어의 의미는 풀고자 하는 문제의 종류 및 범위에 따라 달라질 수 있다. 예를 들어, 단어 '카카오'는 주식 관련 기능으로 분류될 때에는 회사의 명칭으로서의 의미를 가지지만, 교통 정보 관련 기능으로 분류될 때에는 장소 명칭으로서의 의미를 가지며, 음식 관련 기능으로 분류될 때에는 식재료 명칭으로서의 의미를 가진다.

GloVe로 얻어진 단어의 의미 벡터는 여러 의미들을 모두 포괄적으로 나타내고 있으므로 특정 상황에 맞는 의미를 더욱

강화하기 위해 별도의 Local word vector를 구축한다. 해당 벡터는 네트워크 모델 훈련 시 함께 훈련된다.

2) Entity tag vector

사용자의 발화 요청이 짧아서 발화 자체로부터 의미 있는 정보를 얻을 수 없을 경우, 대신 발화에 속하는 엔티티(entity)의 태그(tag) 정보를 이용하여 문장 분류를 진행할 수밖에 없다. 예를 들어, 발화 '아이유 좋은날'의 경우, '아이유'는 인물명이고, '좋은날'이 곡명이라는 사실을 알아야 해당 발화가 음악과 관련된 요청이라는 것을 알 수 있다.

3) POS vector

각 형태소의 품사 정보 또한 사용된다.

문장 훈련을 위해 사용된 네트워크는 EMNLP(2014) 논문⁶에서 제안된 방법으로 구성했다. [그림 1]은 해당 네트워크의 구조를 간략히 보여준다.

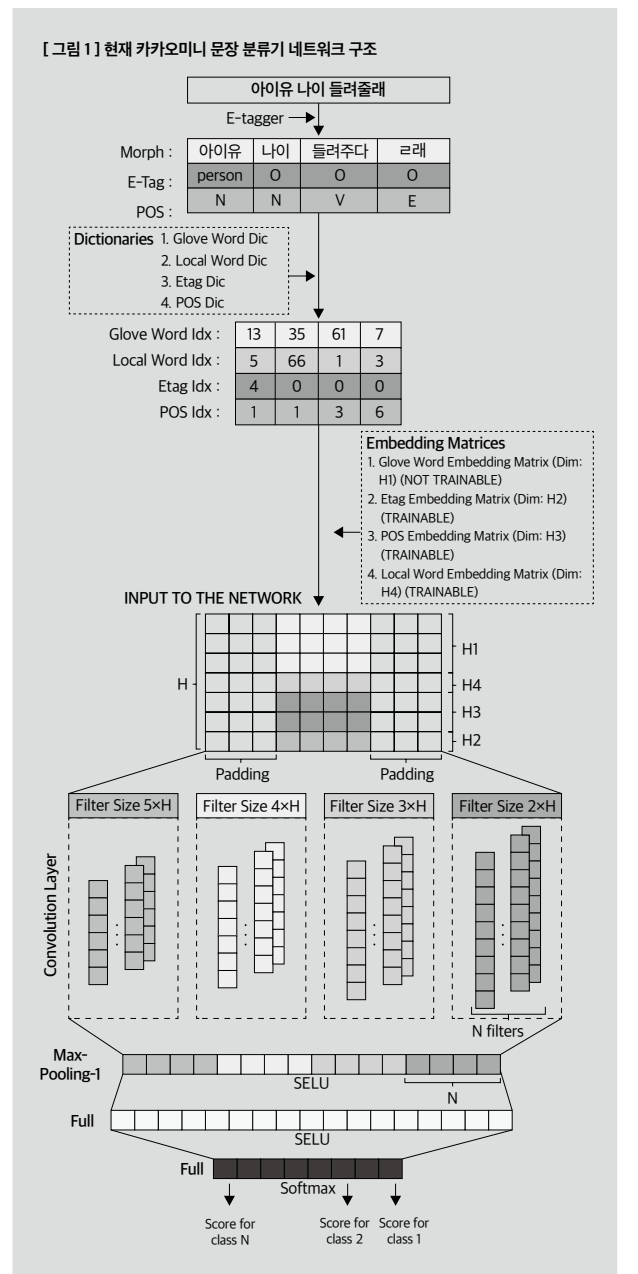
그 밖에 본문에서 자세히 설명되지 않은 내용으로 포지셔널 인코딩(positional encoding)⁷ 및 SELU(scaled exponential linear unit)⁸이 사용되었다.

개선 방향

최근 하나의 문장에 다종의 인텐트를 포함하고 있는 발화들을 처리하는 방법이 구현되어 카카오미니에 적용되었다. 이런 다중 인텐트들의 조합을 가능하게 하고 처리 가능한 조합을 점점 늘려가면서 기존 기능에 악영향을 끼치지 않게 하는 것은 문장 분류 모듈에 있어 굉장히 중요한 도전이 될 것이다.

또한, 인공 신경망을 실제로 사용하기 위해서는 각종 신경망 패러미터(parameter) 튜닝에 많은 시간을 투자하여야 한다. 현재 자동으로 최적의 패러미터를 찾기 위한 작업⁹이 수행되고 있다.

이 작업이 더 발전되면 자동으로 인공신경망 네트워크 구조를 구축하는 시스템¹⁰으로의 발전도 기대할 수 있을 것이다.



^{*1} 논문 | Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE (Vol. 86, pp. 2278-2324). ^{*2} 참고 | 이수경, 이주진, 임성빈. (2018). Brain's Pick - 단어 간 유사도 파악 방법. 카카오 AI 리포트 (Vol. 10). ^{*3} 논문 | Pennington, J., Socher, R., Manning, C. (2014). Glove: Global Vectors for Word Representation. Proceedings of the EMNLP (pp. 1532-1543). ^{*4} 논문 | Bojanowski, P., Grave, E., Joulin, A., Mikolov, T. (2017). Enriching Word vectors with Subword Information. Transactions of the Association for Computational Linguistics (Vol. 5, pp. 135-146). ^{*5} 설명 | 일반적으로 대표적인 확률적 언어 모델의 하나로 n개 단어의 연쇄를 확률적으로 표현해 실제로 발생된 문장의 기록을 계산하는 방법(출처: ITA정보통신용어사전)를 의미하지만, 해당 연구에서는 단어를 이루는 캐릭터들의 연쇄라는 의미로 사용됨. ^{*6} 논문 | Yoon, K. (2014). Convolutional Neural Networks for Sentence Classification. Proceedings of the EMNLP (pp. 1746-1751). ^{*7} 논문 | Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. (2017). Attention is All you Need. Proceedings of NIPS. ^{*8} 논문 | Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S. (2017). Self-Normalizing Neural Network. Proceedings of NIPS. ^{*9} 논문 | Bergstra, J. S., Bardenet, R., Bengio, Y., Kegl, B. (2011). Algorithms for Hyper-Parameter Optimization. Proceedings of NIPS. ^{*10} 논문 | Zoph, B., Le, Q. V. (2017). Neural Architecture Search with Reinforcement Learning. arXiv preprint.

수학자는 어떻게 인공지능을 발전시키는가?

제가 머신러닝(machine learning)이란 용어를 처음 접한 건 2013년도 겨울로 기억합니다. 당시 저는 SNS상으로 종종 인사드리던 고등학교 선배들과 운동 모임을 정기적으로 가졌었는데, 민현석 박사는 그 중 한 분이었습니다. 영상처리와 패턴인식 이론을 전혀 모르는 수학과 대학원생이었던 저는 민현석 박사님 덕분에 이 분야의 발전 가능성을 미리 접할 수 있게 되었습니다. 제가 학계에서 벗어나 산업계로 첫걸음을 내딛게 된 데에는 민현석 박사님의 조언이 매우 큰 영향을 주었습니다.

※ 편집자 주: 물체에도 필자의 의도가 담겨 있다고 판단해서 존칭 표현을 그대로 담았습니다.

딥러닝과 수학의 연결 고리

그러던 어느 날 민현석 박사님이 합성곱 연산(convolution operation)의 수학적인 직관에 대해 물어 보셨습니다. 사실 합성곱 연산은 신호처리 분야 뿐만 아니라 편미분방정식, 확률론 등 현대수학의 해석학(analysis) 분야에서 절대적인 위치를 차지하고 있습니다. 제가 박사과정 때 주로 연구했던 분야도 조화해석학(harmonic analysis)과 확률론이었습니다. 두 이론은 합성곱 연산과 푸리에 분석(Fourier analysis) 또는 웨이블릿 분석(wavelet analysis) 사이의 유기적인 관계를 이용해 시그널의 확률적인 패턴을 분석합니다. 민현석 박사님은 장시간에 걸쳐 딥러닝의 등장 배경과 CNN(convolutional neural networks)의 구조를 설명해 주셨고, 2012년도부터 이미지넷(ImageNet)¹ 등의 대회를 통해서 본격적으로 알려지게 된 딥러닝의 성능 발전에 대해서도 알려 주셨습니다.

한 동안 저는 딥러닝 분야에 수학자가 크게 기여할 수 있는 일이 많지 않다고 생각했습니다. 데이터 기반(data-driven)이면서 모델 비의존적(model-free)인 방법론을 추구하는 학문에서는 수학적인 접근보다 오히려 데이터를 잘 축적하기 위한 파이프라인(pipeline) 설계가 더 중요하다고 생각했기 때문입니다. 저의 이러한 방어적(?) 태도에 대해 민현석 박사님은 다음과 같이 답변했습니다.

“지금으로선 나도 잘 모르겠다. 하지만 딥러닝의 가치가 커질수록 수학이 중요해질 날이 분명 올 것이다.”

딥러닝 엔지니어가 수학을 꼭 전공해야 하나?

딥러닝은 수학적으로 어려운 구조는 아닙니다. 행렬 연산과 비선형 함수(activation function)를 여러 층으로 합성해 쌓는 것이 기본 구조입니다. 1차 미분만 계산하는 역전파(back propagation) 알고리즘은 학부 대학생 정도면 쉽게 구현할 수 있고, 고등학생도 어렵지 않게 이해할 수 있습니다. 게다가 자동미분(automatic differentiation) 라이브러리가 포함된 프레임워크가 오픈소스화 되면서부터 구현하는 것 또한 손쉬운 일이 되었습니다. 텐서플로우코리아(TensorFlow Korea)나 파이토치(PyTorch) 한국 사용자 그룹 같은 SNS 그룹에는 본인이 직접 만든 알고리즘을 깃허브(GitHub)로 공개하는 중학생도 더러 있을 정도입니다. 이런 경향 때문인지 “수학을 전공하는 것이 딥러닝 엔지니어에게 도움이 될까요?”라는 질문이 SNS 상의 커뮤니티에 올라오곤 합니다. 갑론을박이 오고 가는 가운데 전문적인 수학 지식이 왜 필요한가에 대해 대중적인 공감을 얻은 의견은 아직까지 없습니다. 오히려 딥러닝 논문의 수식들은 현학적인 장벽처럼 느껴지고 코드로

아이디어를 이해하는 것이 더 효율적이라는 의견도 있습니다.

저 역시 딥러닝을 목적으로 수학을 깊게 공부하는 것은 시간 낭비라고 생각합니다. 2~3년의 시간을 투자하여 수학과와 전공 수업을 공부하기에는 그 범위가 너무 넓고, 수학 외에도 공부해야 할 더 중요한 지식들이 산더미이기 때문입니다. 게다가 딥러닝의 철학은 기존의 수학적 모델링의 성과를 과감히 포기하고 효율적 연산과 데이터 의존적인 엔드-투-엔드(end-to-end)를 추구하는 점에 있습니다. 그래서 저는 모든 딥러닝 연구자가 수학을 깊게 공부해야 한다는 의견에는 선뜻 동의하기 어렵습니다.

그렇다면 반대 논리로 수학 전공자가 딥러닝 연구에 기여할 가능성은 낮을까요? 그렇지 않다고 생각합니다. 모두가 수학을 전공할 필요는 없지만 수학에 능숙한 연구자는 과학, 공학 연구에 언제나 필요하기 마련인데 가령 주어진 문제에 내재된 제약들을 다른 수학적 조건으로 전환하는 방법을 제시할 수 있습니다. 작년 초에 많이 회자되었던 마틴 아조브스키(Martin Arjovsky)의 베셔슈타인(Wasserstein) GAN(Generative Adversarial Network) 논문이 그 중 하나의 사례로 볼 수 있습니다. 마틴은 기존의 머신러닝 논문에서 주로 사용하던 쿨백-라이블러 발산(Kullback-Leibler divergence) 대신 베셔슈타인 거리함수²를 GAN에 적용하는 아이디어를 수학적인 타당성과 함께 제시했습니다. 그 이후로는 연구자들이 베셔슈타인 거리함수를 쿨백-라이블러 발산을 대신하여 새롭게 시도할만한 손실 함수(loss function)로 인식하기 시작했고, 최근엔 오토 인코더(auto-encoder)³나 강화 학습(reinforcement learning)⁴ 연구에도 도입되기 시작했습니다.

주어진 문제에 알맞은 정규화(regularization) 방법을 제시하는 것도 수학을 전공한 연구자가 현업에서 풀어야 할 과제입니다. 대개 정규화는 과적합(overfitting)을 막고 일반화(generalization) 성능에 도움을 주기 위한 도구로 인식됩니다. 그러나 정규화 방법들을 어떤 문제에도 동일하게 적용할 수 있는 건 아닙니다. 예를 들어 데이터가 상대적으로 적거나 노이즈가 적은 상황에선 모델의 복잡도를 제한할 필요가 있으므로 패러미터(parameter)의 크기를 제한하는 가중치 감소(weight decay) 기법이 제 역할을 하는 경우가 많습니다. 반면, 데이터에 노이즈가 크게 내재된 경우에는 확률적 정규화 기법(stochastic regularization technique)⁵들을 사용하는 베이지안 신경망(Bayesian neural network)이 우수한 성능을 낼 때가 있습니다. 요컨대 정규화 기법은 주어진 학습 문제의 최적해 공간(optimal solution space)을 특정하는 역할로 해석할 수 있고, 최적해 공간이 적절하게 상정되었을 때 일반화 성능이 개선되는 것입니다. 이처럼 “어떤 해를 찾을 것인가?”라는 질문에 대해 데이터 집합의 특징을 고려하여 적절한 선행적 지식(prior knowledge)을

글 | 임성빈 leo.brain@kakaobrain.com

수식과 증명을 사랑하던 수학자가 로보틱스와 인공지능의 매력에 빠지게 되었습니다. 카카오브레인에서 수학과 엔지니어 사이의 가교 역할을 담당하는 수리과학자(Mathematical Scientist)로 활동하고 있습니다. 안전하고, 신뢰도 높은 인공지능을 개발해 우리 사회를 보다 윤택하게 만들고 싶습니다.

반영하는 모델을 찾는 것은 수학적 근거를 필요로 합니다.

또한 수학은 지금까지는 시도되지 않은 종류의 데이터에 접근할 수 있는 방법론을 제시할 수 있습니다. 후자의 대표적인 예시는 최근에 주목을 받고 있는 지오메트릭 딥러닝(geometric deep learning, GDL)⁶입니다. 지오메트릭 딥러닝은 2017년 NIPS(Neural Information Processing Systems) 튜토리얼(tutorial)에 소개 될 정도로 그 중요성이 새롭게 인식되는 분야입니다. 지오메트릭 딥러닝은 데이터를 그래프나 다양체(manifold) 같은 비유클리드(non-Euclid) 공간 상의 대상인 경우로 상정하고, 그 위에서의 딥러닝 네트워크를 학습하는 방법을 다룹니다. 지오메트릭 딥러닝 분야에 대한 상세한 내용은 다음 기회에 다루도록 하겠습니다.

수학자들의 도전을 기다리는 인공지능 이론

지금까지는 수학 전공 연구자들이 딥러닝 연구에 기여할 수 있는 내용들에 대해 살펴보았습니다. 그렇다면 학계의 수학자들은 딥러닝을 어떻게 바라보고 있을까요? 종종 “딥러닝은 블랙박스라 수학적으로 이해할 수 없지 않나요?”라는 질문을 듣게 되는 경우가 있지만, 역사적으로 보면 수학자들은 오래전부터 신경망 연구를 시도하였고 어느 정도 성공을 거둔 영역도 있습니다.

[그림 1] 찰스 페퍼먼(Charles Fefferman)⁷



조화해석학은 물리학과 신호처리 이론의 중요한 근간이 되는 학문입니다. 딥러닝 연구도 이 조화해석학을 통해 많은 성질들이 규명되었는데 찰스 페퍼먼(Charles Fefferman, [그림 1])이 그 대표적인 수학자라고 할 수 있습니다. 페퍼먼은 현대 수학의 다방면에서 활약한 천재 수학자인데 신경망 연구 뿐만 아니라 머신러닝 전반에 걸쳐 어려운 문제들을 풀어내고 있습니다. 그 중 하나가 다양체 가설(manifold hypothesis) 문제입니다. 이는 고차원의 데이터를 차원을 축소하여 저차원의 데이터로 표현하는 다양체의 존재성에 대한 가설입니다. 다양체 가설은 머신러닝

연구자들이 대부분의 이미지나 텍스트 데이터에서 참으로 간주합니다. 여러 실험 결과들이 이를 뒷받침하고 있지만⁸, 이를 수학적으로 증명하는 것은 어려운 일입니다. 페퍼먼은 이 다양체 가설이 수학적으로 검증 가능하다는 것을 2016년 Journal of AMS에 게재된 ‘Testing the Manifold Hypothesis’ 논문에서 증명하였습니다.

[그림 2] 스테판 말렛(Stephane Mallat)(좌)⁹과 조안 브루나(Joan Bruna)(우)¹⁰



스테판 말렛(Stephane Mallat, [그림 2])과 조안 브루나(Joan Bruna, [그림 2])는 2012년 ‘Invariant Scattering Convolution Networks’¹¹라는 논문에서 웨이블릿 필터(wavelet filter)를 가진 CNN 구조가 가지는 수학적 성질들을 밝혀낸 적이 있습니다. 구체적으로 ‘CNN은 차원의 저주(curse of dimension)를 극복하고 이미지 데이터에 내재된 불변량(invariant)을 찾아낸다’라는 추측을 수학적으로 참인 명제로 증명해 낸 것입니다. 이 논문의 뒤를 이어서 토마스 와이토스키(Thomas Wiatowski)와 헬멧 볼스키(Helmut Bolcskei)는 일반적인 필터들로 이루어진 CNN도 ‘Invariant Scattering Convolution Networks’ 논문에서 말한 CNN 성질과 같은 성질을 만족함을 증명했습니다.¹² 이 외에도 ‘Mathematics of Deep Learning’¹³ 논문을 보시면 딥러닝 연구에 기여한 수학적 결과들을 개괄적으로 보실 수 있습니다.

앞에서 소개한 수학자들의 연구 결과들도 놀랍지만 인공지능 분야에서의 수학 연구들은 많은 부분에서 현재 진행형입니다. 이 분야에서 수학이 필요한 이유는 재현 가능성(reproducibility)의 보장, 이론적인 예측과 실험 결과의 일치가 필요 조건인 과학 이론을 정당화하는 것이 바로 수학의 역할이기 때문입니다. 물론 ‘모든 과학 이론은 수학적 증명 없이는 이론이 될 수 없다’는 환원주의적인 주장이 항상 옳은 것은 아닙니다. 그러나 인공지능 이론이 그 범주 밖에 있다고 생각하지 않습니다. 다비드 힐베르트(David Hilbert, 1862-1943)의 제자 중에 리하르트 쿠란트(Richard Courant, 1888-1972)라는 수학자가 있었습니다. 쿠란트는 물리학과 현대 응용수학의 발전에 수많은 기여를 했던 사람인데, 생전에 다음과 같은 말을 남긴 적이 있습니다.

“실험적 증거(empirical evidence)는 결코 수학적 존재성(existence)을 의미하지 않으며, 엄밀한 증명에 대한 수학자들의 요구를 쓸데없는 것으로 여기게 하지도 않는다. 현상에 대한 수학적 설명이 의미를 가지게 만드는 건 오로지 수학적인 증명을 통해 가능한 것이다.”¹⁴

2017년도 여름에 열린 제주 머신러닝 캠프에서 뉴욕대학교(New York University, NYU)의 조경현 교수님이 “딥러닝이 정말 블랙박스인가?”라고 질문을 던진 적이 있습니다. 현재는 딥러닝을 포함해 인공지능 분야의 연구가 실험적인 결과들에 의해 대부분 진행이 되고 있습니다만, 언젠가 상당수의 명제들이 수학적으로 증명된 명제들로 논증되는 시점이 올 것이라 생각합니다. 그리고 그렇게 이루어진 인공지능 연구가 우리의 삶을 더 윤택하게 만들 수 있다고 저는 확신합니다. 어쩌면 오랜 시간이 필요한 작업이 될 수 있을 것입니다. 그러나 힐베르트의 명언을 인용하면 우리는 알아야만 하며 언젠가는 알게 될 것입니다.

글을 맺으면서

예전에 모 대학의 수학과 교수님께서 저에게 “어떤 수학 전공자를 인공지능 업계에서 선호합니까?”라고 물어 보신 적이 있습니다. 그 분께서는 “어떤 전공이 중요한지에 대한 답을 기대하셨지만, 저는 “특정 전공보다 소통 능력이 더 중요합니다”라고 답한 바 있습니다.

저는 수학자들의 연구가 비약적인 성공을 거둔 후에도 공학을 포함한 비전공자들의 직관적인 이해를 돕거나 성능 향상에 도움이 되려면 수학자들의 더 많은 노력이 필요하다고 생각합니다. 왜냐하면 용어가 지나치게 전문적인데다 수학자들의 논증이 실제적인 예제로 설명하기가 어려울 정도로 비전공자의 시각에선 추상적이기 때문입니다. 사실 수학이 추상화를 추구하게 된 원인은 지난 2세기 동안 과학의 발전에 성공적으로 공헌했던 것이 큼니다. 이는 학문적 가치관으로서의 옳은 방향이었지만, 공학적 연구에 도움이 되려면 이제는 소통의 관점에서 접근해야 합니다. 저는 최근에 카카오브레인에 합류한 최성준 박사님과 로보틱스 분야에서의 머신러닝 활용을 연구하고 있습니다. 막연하게 떠오른 아이디어를 수학적 언어로 토론하고 코드로 구현하여 실험합니다. 그렇게 구체화되는 연구 결과들을 보면서 매일 큰 보람을 느끼고 있습니다. 그 과정에서 공학자와 수학자 간의 협업에 있어 위에서 말한 소통이 매우 중요하다는 점을 깨닫고 있습니다.

제가 생각하는 STEM¹⁵업계에서의 성공적인 조직 문화 요건들 중 빼놓을 수 없는 것이 ‘소통’ 과 ‘다양성’¹⁶입니다. 소통과 다양성이 왜 필요할까요? ‘다양성과 복잡성(Diversity and Complexity)’의 저자인 스콧 페이지(Scott Page)는 복잡한

과학적 문제의 해결은 단순히 한 분야의 전문성에 의존하기 보다는 기존의 접근 방식과 새로운 방식을 접목하고 비틀어보는 ‘발상의 전환’에서 나온다고 설명합니다.¹⁷ 복잡한 문제일수록 변수와 구조가 다차원적이기 때문입니다. 새로운 발상이 통찰력 있는 아이디어가 되려면 문제를 바라보는 다양한 시각들이 얼마나 서로 소통하고 적용되는지에 따라 결정됩니다. 페이지 교수는 이러한 소통과 다양성의 메커니즘을 머신러닝 알고리즘 중 하나인 랜덤 포레스트(random forest)¹⁸에 비유했습니다. 복잡계 문제를 푸는 알고리즘과 팀 조직 구성이 비슷한 면모를 가진다는 점이 흥미롭습니다.

인공지능 연구는 아직 미완성이며 그 가운데 해결해야 할 많은 난제들이 쌓여 있습니다. 앞으로 다양한 수학 전공자들이 인공지능 연구 분야에 뛰어들어 기술 발전에 공헌할 날이 오기를 기대해 봅니다.

¹ 설명 | 2010년부터 매년 개최되는 대회로 물체 인식, 물체 검출, 동영상 물체 검출, 장면 분류, 영상 분할 등 인공지능 및 딥러닝 관련 문제를 해결하는 경연대회로 구글, 마이크로소프트, 퀄컴 등 세계적 IT 기업이 출전하는 대회이다. ² 참고 | 두 거리함수에 대한 구체적인 비교는 Wasserstein GAN 논문 및 링크 자료 (<https://www.slideshare.net/ssuser7e10e4/wasserstein-gan-i>) ³ 참고 | Wasserstein Auto-Encoders (<https://arxiv.org/abs/1711.01558>) ⁴ 참고 | On Wasserstein Reinforcement Learning and the Fokker-Planck equation (<https://arxiv.org/abs/1712.07185>) ⁵ 설명 | 페라미터에 사전(prior) 확률분포를 부여해서 노이즈의 성능감소를 감쇄하는 페라미터를 학습하는 기법. 드롭아웃(dropout)이 대표적인 방법이다. ⁶ 참고 | <http://geometricdeeplearning.com/> ⁷ 참고 | <https://tecnico.ulisboa.pt/en/events/symposium-interpolation-and-approximation-in-several-variables/> ⁸ 논문 | Ian Goodfellow et al. Deep Learning, p.161 ⁹ 참고 | <https://www.di.ens.fr/~mallat/> ¹⁰ 참고 | <https://cims.nyu.edu/~bruna/> ¹¹ 논문 | Invariant Scattering Convolution Networks (<https://arxiv.org/abs/1203.1513>) ¹² 논문 | A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction (<https://arxiv.org/abs/1512.06293>) ¹³ 논문 | Mathematics of Deep Learning (<https://arxiv.org/abs/1712.04741>) ¹⁴ 원문 | Empirical evidence can never establish mathematical existence nor can the mathematician's demand for existence be dismissed by the physicist as useless rigor. Only a mathematical existence proof can ensure that the mathematical description of a physical phenomenon is meaningful. ¹⁵ 설명 | Science, Technology, Engineering, Mathematics의 약자 ¹⁶ 참고 | 보통 다양성(diversity)은 성(gender), 인종(race), 사회경제(socioeconomic)적인 배경 등의 문화적 차이를 의미하지만 본 글에서는 전공 별 다양성에 초점을 두었습니다. ¹⁷ 참고 | <https://steemit.com/writing/@sungyu1223/the-diversity-bonus-how-great-teams-pay-off-in-the-knowledge-economy-our-compelling-interests> ¹⁸ 설명 | 랜덤 포레스트는 부트스트랩(bootstrap)시킨 훈련데이터에 랜덤하게 선택된 서로 다른 변수들을 포함하는 나무 모형(tree model)으로 적합시키는 앙상블(ensemble) 기법으로, 성능을 높이면서 동시에 과적합을 방지합니다.

딥페이크와 리벤지 포르노

이슈 | 딥페이크와 페이크앱

드라마 | 블랙리더 ‘닥치고 춤춰라’

2017년 4월 캐나다 몬트리올 기반 인공지능 스타트업인

라이어버드(Lyrebird.ai)는 인간의 목소리를 흉내 내는 음성

합성(speech synthesis) 기술을 선보였다. 라이어버드는 약

60초 분량의 음성 샘플 데이터만으로 화자의 말씨나 억양을 거의

그대로 재현했다. 목소리 재현에 20분 가량의 샘플을 필요로 하는

어도비(Adobe)의 ‘프로젝트 보코(Project VoCo)’²와 비교하면, 적은

데이터로도 비슷한 성능을 낸다는 점은 라이어버드의 경쟁력 있는

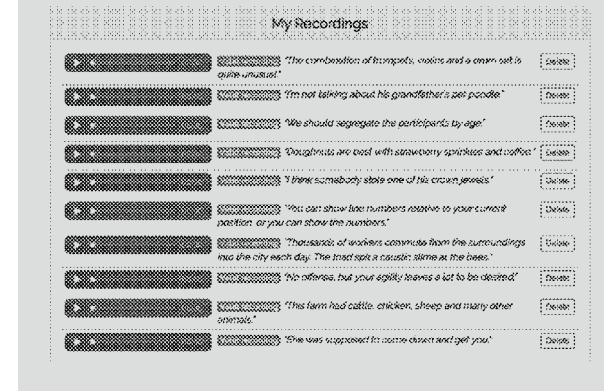
요소라고 할 수 있다.

컴퓨터로 음성과 영상을 생산하는 시대

라이어버드는 이런 획기적인 기술을 활용해 버락 오바마(Barack Obama) 전(前) 미국 대통령, 도널드 트럼프(Donald Trump) 미국 대통령, 힐러리 클린턴(Hillary Clinton) 전(前) 미국 국무장관 등과 같은 유명인의 목소리를 흉내 낸 데모 버전³을 공개함으로써 자사 기술력을 널리 알렸다.

그렇다면 인공지능은 유명인의 목소리가 아닌, 일반인의 목소리도 흉내 낼 수 있을까? 라이어버드 홈페이지에서 내 목소리를 통해 직접 테스트해 볼 수 있다. [그림1]에서 보듯이 회사 측이 제시하는 문장을 최소 30개 이상 읽은 음성 파일을 서버에 올리면, 라이어버드 엔진은 이 파일을 학습해 그 즉시 사용자의 목소리를 재현해낸다(학습에는 수 분이 걸린다). 0.5초 안에 1,000개의 문장을 생성할 수 있는 GPU(Graphics Processing Unit) 클러스터를 갖춘 덕분이다. 더 많은 문장을 읽을수록 즉 학습 데이터 셋이 많아질수록 실제 사용자와 더욱 유사한 목소리를 흉내 낸다. 단, 제약 조건이 하나 있다. 미국식 영어 발음에 최적화되어 있다는 점이다.

[그림 1] 라이어버드에서 음성합성에 필요한 데이터를 녹음한 예시¹



마찬가지로 영상도 손쉽게 만들어 낼 수 있다. 단적인 예로, 얼굴 합성 기술을 활용하면 단순한 조작만으로도 자신의 얼굴에 특수 효과를 입힐 수 있다. 또한, 얼굴을 나이 들어 보이게 만들거나 남자를 여자로 바꾸는 것도 가능하다. 자글한 이마 주름을 지우거나 헤어 스타일도 바꿀 수 있다. iOS 앱스토어의 ‘페이스앱(FaceApp: Neural Face Magic)’⁴은 이 마법과 같은 기능을 모두 가능하게 하는 앱이다.

지난 2016년 미국 스탠퍼드(Stanford) 대학교와 독일 에르랑겐-뉘른베르크(Erlangen-Nuremberg) 대학교, 독일 막스플랑크 연구소(Max Plank Institute)로 구성된 연구팀은 ‘Face2Face’라는 얼굴 치환 프로그램을 선보였다.⁵ 이 기술은 사용자(source)의 입 모양과 표정을 영상 속 인물(target)에 입히는 것으로 사용자가 원하는 말이나 표정을 영상 속 인물이 직접 하는 듯한 영상으로 실시간 변환하여 보여준다.



기술적 배경

음성합성은 오래전부터 존재한 기술이지만 생성된 음성의 품질은 확연히 낮았다. 한편, 컴퓨터로 비디오표를 만들어내는 작업은 대규모 자본을 갖춘 할리우드 제작사나 최신 기술을 연구하는 연구소에서나 가능한 것으로 여겨질 정도로 막대한 자본력과 기술력을 필요로 했다. 하지만 데이터 분석에 주로 사용했던 딥러닝 기술 덕분에 새로운 음성과 영상 콘텐츠를 보다 손쉽게 만들 수 있게 됐다. 물론 그래픽 하드웨어와 소프트웨어의 발전, 새로운 비디오 캡처 기술 또한 이러한 기술적 변화를 끌어내는 데 영향을 미쳤다.

음성합성의 시초

음성합성에 관한 아이디어를 처음 고안하게 된 시기는 전기 신호 처리가 나온 1989년 이후 부터다. 그러다 벨 연구소(Bell Labs)에서 보코더(vocoder)⁶를 개발한 1930년대 이후 실제 기술 구현도 가능해졌다. 보코더의 개념은 간단하다. 음성에서 중요한 부분만 추출하고, 이를 활용하여 반대로 음성 신호를 재구성한다. 다시 말해 음성 신호의 대역폭(특정 시간 내에 보낼 수 있는 정보량) 또는 주파수 대역을 줄이는 과정이다.

여기에 활용된 접근 방식은 바로 생성 모델(generative model)이라는 기술이다. 생성 모델은 샘플 데이터를 통해 생성하려는 대상의 특성(확률분포)을 학습한 뒤, 이와 유사한 데이터를 생성해낸다. 무엇을 생성하는지는 적용하는 대상에 따라 이미지, 소리, 텍스트 등이 될 수 있다.

위에 언급된 라이버드와 페이스앱은 신경망을 활용한 생성 모델(deep generative networks)을 구축한다. 기술적으로는

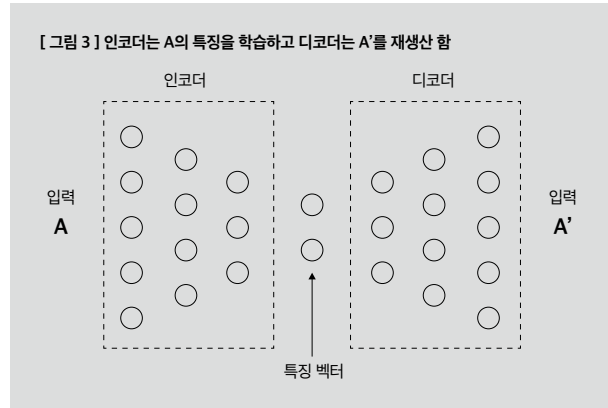
글 | 이수경 samantha.lee@kakaobrain.com

2016년 3월 알파고와 이세돌 9단이 펼치는 세기의 대결을 두 눈으로 목도한 이후 인공지능을 제대로 공부해 봐야겠다고 결심했습니다. 인공지능 본진이자 연구소인 카카오브레인으로 걸어 들어온 이유입니다. 인공지능 기술과 이로 인해 바뀔 미래 사회를 다루는 글을 통해 사람들과 소통하고 싶습니다.

감수 | 홍상훈 sanghoon.hong@kakaobrain.com

카카오브레인에서 새로운 아이디어를 고민하고 이를 직접 구현하는 리서치 엔지니어로 활약 중입니다. 학부 시절부터 인공지능, 뇌 연구에 관심이 많았습니다. 딥러닝의 잠재력에 매일 놀라고 있지만, 완전한 인공지능이 될 때까지는 아직 갈 길이 멀다고 생각합니다. 카카오브레인의 뛰어난 사람들과 함께 AI 기술 발전에 이바지함으로써 인류 사회에 보탬이 되고자 합니다.

인코더(encoder)와 디코더(decoder)에 깊은 층을 쌓아 거대한 신경망 구조를 갖추는 것으로 보면 된다. 인코더는 방대한 데이터를 통해 다양한 얼굴 또는 목소리를 표현할 수 있는 주요 특징(feature)을 학습하고, 디코더는 주요 특징 정보를 기반으로 원본과 유사한 데이터를 생성하는 방법을 배운다. 이것은 충분한 데이터가 있다면 가능한 일이다. 학습 과정에 대한 자세한 내용은 뒤에서 확인할 수 있다.



선봉적인 인기를 끈 페이크앱(FakeApp)

지난해 딥러닝 기반 얼굴 합성 기술인 딥페이크(deepfake)를 손쉽게 적용할 수 있는 데스크톱 앱(app) '페이크앱(FakeApp)'이 선봉적인 인기를 끌면서, 일반인들도 손쉽게 비디오 영상 속 얼굴을 다른 얼굴로 합성할 수 있게 됐다. 예를 들면, 아나운서 방송 비디오에 미키마우스 얼굴을 합성해 마치 미키마우스가 뉴스를 전달하는 듯한 영상을 생산할 수 있다.

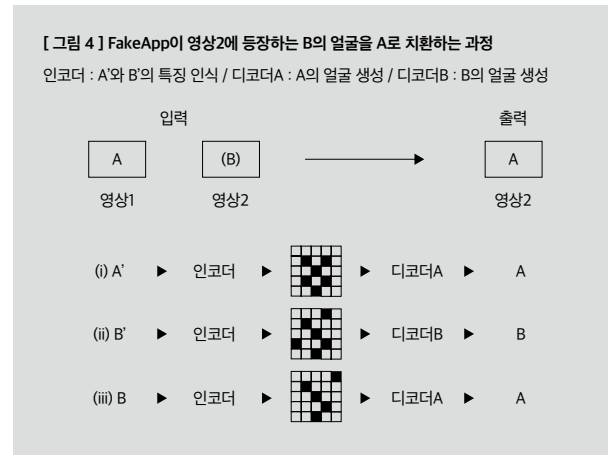
물론 합성된 영상을 구별하지 못할 정도의 완벽한 기술력을 갖추지는 못했다. 얼굴 윤곽 부분에 남아 있는 합성의 흔적 또한 쉽게 찾아볼 수 있다. 이런 기술적 한계에도 불구하고 사람들의 호기심을 자극한 페이크앱은 지난 1월 레드잇(reddit)에 소개된 이후 12만 회의 다운로드를 기록했다.

페이크앱을 만든 익명의 소프트웨어 개발자 'N'은 뉴욕 타임스(New York Times)와의 인터뷰에서 창조적인 실험을 위해 해당 앱을 개발했다고 말했다.⁷ 그는 누구나 한 번쯤 가지고 놀고 싶어 했을 법한 기술을 일반인들도 쉽게 사용할 수 있도록 패키지 형태로 내놓은 게 페이크앱이라고 설명했다.

페이크앱은 구글(Google)이 지난 2015년 공개한 머신러닝 도구인 텐서플로(TensorFlow)를 활용한다. 텐서플로는 시행 착오(trial and error)를 거치면서 이미지를 잘 인식할 수 있도록 학습한다. 데이터 학습에 필요한 것은 좋은 성능의 컴퓨터다. 데이터 처리 능력이 뛰어날수록 수 주간 걸리는 학습 시간을 단 몇

시간으로 크게 줄일 수 있기 때문이다. 또는 구글 클라우드 플랫폼의 서버를 빌리는 방식도 고려해 볼 수 있다.⁷

서버를 세팅하고 페이크앱을 올려 놓았다면 이제 남은 일은 데이터를 모으는 것이다. 실험결과, 대략 300장~2,000장 정도의 이미지 소스를 갖췄을 때 좋은 결과물을 얻을 수 있는 것으로 알려졌다. 그리고 여러 각도에서 촬영한 것보다 일정한 각도에서 촬영한 영상이 더 우수한 학습 결과를 낸다. 아울러 짧은 분량의 비디오 클립이 조작하기 더 쉽고 얼굴형이 서로 비슷한 영상일수록 학습 효과도 좋다. 문제는 학습에 활용할 이미지 데이터 셋을 확보하기가 쉽지 않다는 점이다. 그래서 특징 학습에 한정하여 자신의 얼굴을 촬영한 사진이 부족하다면 자신과 느낌이 비슷한 유명 배우의 사진으로 학습 데이터를 모으는 것도 팁이다.



개발자가 구축한 딥페이크에 데이터를 넣으면 얼굴이 합성된 영상을 얻을 수 있다.⁸ 영상2에 등장하는 B의 몸에 A의 얼굴을 덮는 과정은 [그림 4]에 자세히 설명되어 있다. 여기서 중요한 점은 인코더에 입력하는 데이터는 임의로 변형한 A'와 B'라는 점이다. 얼굴 일부를 찌그러뜨리거나 늘린 것이다. 이는 원본과 유사한 영상으로부터 원본을 생성하는 방식을 배우게 하기 위한 조치다.

[그림 4] (i)에서 인코더는 A'의 특징을 학습하고, 디코더A는 원본 A를 생성하는 법을 학습한다. (ii)에서 인코더는 B'의 특징을 학습하고, 디코더B는 원본 B를 생성하는 법을 학습한다. (i)과 (ii)를 반복하면 인코더는 A'와 B', 두 사람 얼굴에서 공통적인 얼굴 특징을 뽑아내고, 디코더A와 디코더B는 각각 A와 B를 생성하는 법을 습득한다.

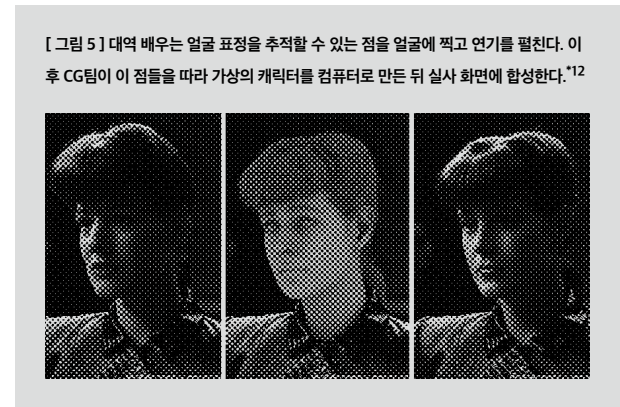
충분한 학습을 거쳐 마지막 단계 (iii)에서는 영상2를 넣으면 인코더-디코더A의 과정을 거쳐 영상2 속 얼굴(B)이 A로 바뀐다. 모든 이미지 프레임에서 이 과정을 반복하고, 얼굴이 부드럽게 움직이도록 하는 휴리스틱 기법을 적용하는 등의 후처리를 거쳐 영상의 품질을 높일 수 있다.

딥페이크가 잠재력을 발휘할 분야는?

앞서 일반인들도 페이크앱을 내려받으면 영상 속 얼굴을 손쉽게 변환할 수 있다고 설명했다. 하지만 기본적인 개발 지식과 충분한 학습 데이터 셋을 갖춰야 한다는 점에서 마냥 쉬운 작업은 아니다. 라이어버드의 공동 창업자인 알렉산드르 드 브레비슨(Alexandre de Brébisson)은 “비디오와 같은 고차원 공간은 변동성이 높기에 더 높은 기술력이 뒷받침되어야 한다. 이제 첫걸음을 떼었다는 점에서 사실 이 기술에는 보완할 부분이 더 많다”고 말했다.

그렇지만 현재 기술 발전 속도를 본다면 지금처럼 스마트폰 앱스토어에 접속해 앱을 내려받고, 해당 앱을 실행하기까지 오랜 시간이 걸리지 않을 것으로 예상된다. 페이크앱 개발자 'N'은 딥러닝 알고리즘이 스마트폰이나 데스크톱 애플리케이션에 활용되는 것 뿐만 아니라, 많은 전자 제품에 적용되는 '기본'적인 기능 중 하나로 자리하게 될 것이라고 내다봤다. 실제로 페이크앱과 같은 딥페이크 기술이 고도화된다면 촬영이나 녹음과 같은 기록 행위를 하지 않고도 실제와 같은 동영상과 오디오 파일을 생성할 수 있을 것으로 보인다.

이런 기술이 가장 활발하게 쓰일 분야는 바로 영화 속 CG(컴퓨터 그래픽)이다. 컴퓨터 그래픽 기술자가 아무 것도 없는 상태에서 인간처럼 움직이는 무언가를 만들지는 못한다. 이런 점을 극복하기 위해 배우에게 수트를 입혀 배우의 움직임을 포착하는 모션 캡처(motion capture) 기술이 고안됐다. 여기에 더 나아가 인간의 구체적인 몸 동작은 물론, 표정과 감정까지 담아내는 이모션 캡처(emotion capture) 기술도 있다.⁹ 이런 점에서 알 수 있듯이 얼굴을 새로 만들어내는 기술은 보편화돼 있다.¹⁰ 문제는 모션 캡처 장비, 모션 컨트롤 카메라를 구매하고 이 장비를 다루는 전문 인력을 고용하는 비용이 높다는 점이다.¹¹ 또한 후처리 작업에도 상당한 시간을 쏟아 부어야 한다.



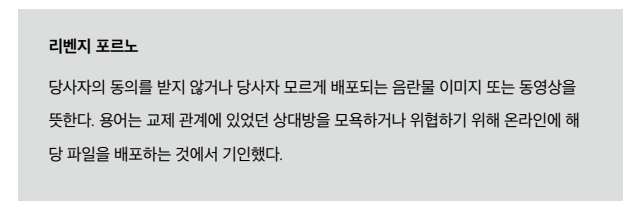
예를 들어보자. 영화 '블레이드 러너 2049(Blade Runner 2049)'의 VFX(visual effects, 비주얼 이펙트)팀은 30년 전 원작 '블레이드 러너'에서 레이첼 역을 맡은 배우 손 영(Sean Young)을 재현하고자

CG 기술을 활용했다. 문제는 몇 분 남짓한 장면을 제작하는 데 무려 10개월이나 되는 기간이 소요되었다는 점이다. 그리고 긴 제작 기간은 영화 제작 비용(인건비) 상승으로 이어질 공간이 크다.

물론 딥러닝이 모든 CG 기술을 대체하지는 못할 것이다. '블레이드 러너'에서 처럼 적어도 수십년 전의 영화를 리메이크할 때 과거 출연한 배우의 젊은 시절 혹은 생전 모습을 되살리는 데 특기를 발휘하리라 기대해 볼 수 있다. 물론 전문 촬영 장비나 전문 인력을 갖추지 않고도 저예산 영화에서 수준급의 CG 효과를 적용할 수도 있다. 그저 재현해내려는 캐릭터와 비슷한 체격과 얼굴을 갖춘 대역 배우만 있으면 된다.

위조된 포르노가 판치다

이처럼 누구나 손쉽게 페이크앱에 사용된 딥페이크를 사용할 수 있게 된다면 경제적인 부가가치를 얻을 수 있는 동시에 악용 가능성도 커진다. 누구보다 이미지를 중요시하는 정치인이나 배우, 아나운서 등 유명인의 얼굴을 입력한 조작된 포르노나 리벤지 포르노(revenge porn)를 양산하는 것이 단적인 예다. 이런 이유로 최근에는 딥러닝으로 조작된 포르노를 만드는 행위 자체를 딥페이크라 부르기도 한다.



레딧에 딥페이크 채널이 활성화되고 많은 사람들 사이에서 이슈화되면서 일부 사람들이 유명인의 거짓 포르노를 생산하는 등의 딥러닝을 악용하는 문제가 생겼다. 해외 IT 매체 마더보드(Motherboard)는 레딧의 커뮤니티가 유명인들의 얼굴을 합성한 포르노 영상을 생성하는 일을 벌여왔다고 최초 보도¹³하기도 했다. 사진이 저장된 스마트폰이나 클라우드를 해킹해 사진을 유출한 사건인 패프닝(Fapping)처럼 연출하면서 말이다. 스칼렛 요한슨(Scarlett Johansson), 갈 가돏(Gal Gadot), 데이지 리들리(Daisy Ridley), 제시카 알바(Jessica Alba), 테일러 스위프트(Taylor Swift), 엠마 왓슨(Emma Watson)과 같은 할리우드 여배우들이 희생양이 됐다.

미셸 오바마(Michelle Obama) 미국 전(前) 영부인도 이 사건에서 예외는 아니었다. 이는 특정 정치인 후보를 비방할 목적으로 위조된 포르노 영상이 제작된 것으로 보인다. 누가 봐도 미셸 오바마로 보이는 한 여성이 검은색 브래지어와 목이 깊게 파인 상의를 입은 채 카메라를 응시하며 이상한 미소를 지어 보인다.

그리고 미셸의 도플갱어는 상의를 탈의한 채 스트립쇼까지 벌인다. 물론 이 영상은 페이크 비디오(fake videos)다. 합성한 흔적을 보면 누구라도 영상이 위조된 것이라는 걸 알 수 있지만 이 역시 자세히 보지 않으면 진짜라 착각할 여지가 크다.¹⁴

이처럼 악용의 우려가 높다고 판단한 트위터와 같은 사이트는 딥페이크로 만든 영상의 유통을 차단했다. 레딧은 10만 명 규모에 달하는 딥페이크 커뮤니티를 폐쇄하기에 이른다. 이런 조치를 두고 일부는 딥페이크가 가진 잠재적인 위험성을 과대 평가하는 미디어를 비난했다. 플랫폼 사들이 거짓 포르노를 엄중히 단속할 것을 염려해 해당 영상물을 다른 플랫폼으로 실어 나르는 부류도 있었다. 극히 일부만이 딥페이크 행위가 양심에 어긋난다고 말했다는 뿐이다.

진실과 거짓 사이를 아슬하게 걷는 현대인들

드라마 '블랙미러 : 닥치고 춤춰라(Shut up and Dance)'의 주인공 케니는 본인의 은밀한 사생활을 담은 동영상이 외부로 유출되는 사건을 겪는다. 케니의 노트북에 접근한 해커는 웹캠으로 촬영한 동영상을 지인에게 유포하겠다고 케니를 협박한다. 패닉에 빠진 케니는 해커가 지시하는 대로 행동한다. 하지만 수많은 랜섬웨어(ransomware) 사태에서 보듯이 해커는 케니와의 약속을 저버리고 그의 사생활 영상을 외부로 유출한다.¹⁵ 이것으로 케니의 사회 관계가 파괴됐음은 자명한 일이다.



딥페이크와 같은 기술로 인해 은밀한 사생활을 찍은 동영상이 실제로 존재하지 않더라도 패프닝은 발생할 수 있다. 앞서 언급한 대로 비단 유명인들에게만 일어날 수 있는 것은 아니다. 일반인들도 희생양이 될 수 있으며 명예훼손, 명의 도용, 사실 왜곡 등 사회 문제를 유발할 위험도 높다.

이러한 패프닝은 다량의 데이터만 있다면 충분히 가능한 일이다. 온라인에 넘쳐나는 셀피(selfie) 덕분이다. 지난 2016년 구글 포토(Google Photos)에는 240억개의 셀피가 업로드됐다.¹⁷ 같은 해 인스타그램(Instagram)에서 셀피 해시태그(#selfie)가

달린 사진만 2,900만 개가 업로드됐다. 셀피를 뜻하는 각종 언어로 작성된 태그까지 포함한다면 그 수는 더 많을 것이다.¹⁸ 때문에 악의를 가진 누군가가 포르노 영상에 피해자의 사진을 합성해 유포하기가 더 쉬워졌다. 딥페이크의 출현 이전에 1,000만 명 중 25명(미국 기준)¹⁹이 위조된 포르노 사진으로 협박을 받았다는 것을 감안한다면 앞으로 이 수치는 계속 늘어날 수 있다.

유명인은 동의 없이 자신의 이미지를 상업적 혹은 불법적인 목적으로 사용했다는 것을 이유로 소송을 제기할 수 있다. 예를 들어, 갠 가돗의 허가 없이 그녀의 사진을 스트립 클럽 홍보에 사용하면 명백한 불법이다. 일반인들은 '정보통신망 이용촉진 및 정보보호 등을 위한 법률'상 명예훼손 등 권리를 침해받았다고 주장할 수 있다. 문제는 해외 사이트를 통해 유입되는 경우 또는 개인이 SNS 등을 통해 공유하는 경우다. 나도 모르게 해외 사이트에 내 얼굴이 합성된 위조된 포르노 사진이 유포된다는 사실은 알기도 힘들거니와 알더라도 처벌 또한 어렵다.

현재 관련 기술 연구에 가속도가 붙고 있는 만큼 미래에는 위조된 비디오와 오디오를 판별하는 게 더욱 중요한 이슈로 부상할 것은 너무나도 자명한 일이다. 메릴랜드 대학교(University of Maryland)의 컴퓨터 과학자인 젠 골벡(Jen Golbeck)은 "유튜브와 같은 글로벌 콘텐츠 플랫폼은 포르노 감지 기능을 탑재할 필요가 있다"고 역설한다. 이런 문제를 인식한 과학기술정보통신부도 합성 사진을 자동으로 판별해 악의적인 유포를 효과적으로 차단할 수 있는 인공지능을 개발하는 연구개발 대회²⁰를 진행하고 있다.

그렇지만 소셜 미디어 플랫폼 사가 위조된 정보를 판별해 유통을 차단하는 알고리즘을 고도화해 나가는 것만이 원천적인 해결책은 아니다. 위와 상관없이 위조된 콘텐츠가 지금도 수많은 '좋아요'를 받고 전 세계로 퍼져나가고 있다는 점을 감안한다면, 조작의 행위가 줄어들기는커녕 계속 늘어날 것이다. 이 상황은 특정인만이 아닌 우리 모두의 이야기가 될 수 있다. 근본적인 대책이 마련되어야 하는 이유다.

서던 캘리포니아 대학교(University of Southern California) 컴퓨터과학과의 하오 리(Hao Li) 조교수는 "선택의 여지가 없다"고 말한다. 이런 인공지능 기술로 조작된 콘텐츠를 만드는 행위가 차세대 커뮤니케이션 방식이라면, 다른 이를 협박하거나 나쁜 일을 도모하려는 목적으로 기술을 악용하려는 이들의 행위 또한 자연스럽게 일어난다는 이유에서다. 그는 "사용자를 끊임없이 교육해야 하는 이유가 바로 이런 우려를 조금이라도 불식하기 위한 것"이라고 덧붙였다.

¹ 참조 | <https://lyrebird.ai> ² 참조 | <http://bit.ly/2zBstpl> ³ 참조 | <https://lyrebird.ai/demo/> ⁴ 참고 | <https://itunes.apple.com/us/app/faceapp-free-neural-face-transformations/id1180884341?mt=8> ⁵ 논문 | https://web.stanford.edu/~zollhofer/papers/CVPR2016_Face2Face/paper.pdf ⁶ 참고 | 도서 '앨런 튜링의 이미테이션 게임' ⁷ 참고 | <https://www.nytimes.com/2018/03/04/technology/fake-videos-deepfakes.html> ⁸ 참고 | <https://hackernoon.com/exploring-deepfakes-20c9947c22d9> ⁹ 참조 | 도서 '입체영화', 도서 '미장센' ¹⁰ 참조 | <https://www.youtube.com/watch?v=MMa2oTlwMIs> ¹¹ 참조 | <http://www.itfind.or.kr/WZIN/jugidong/1136/113605.htm> ¹² 참조 | <https://www.youtube.com/watch?v=fV34mT5m0bM> ¹³ 참고 | https://motherboard.vice.com/en_us/article/gdydydm/gal-gadot-fake-ai-porn ¹⁴ 참고 | 실제 영상을 보면 현재 딥페이크 기술 수준은 상당히 자연스럽다는 것을 알 수 있다. <https://www.youtube.com/watch?v=ttGUwftYvg> ¹⁵ 참조 | <http://www.ciokorea.com/news/30480> ¹⁶ 참고 | <http://www.geekgirlauthority.com/black-mirror-recap-s3-e3-shut-up-and-dance/> ¹⁷ 참조 | <https://blog.google/products/photos/google-photos-one-year-200-million/> ¹⁸ 참조 | <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5445188/> ¹⁹ 참조 | https://datasociety.net/pubs/oh/Nonconsensual_Image_Sharing_2016.pdf ²⁰ 참조 | <https://www.gov.kr/portal/ntnadmNews/1367056?pageIndex=6&hideurl=N>

2018년 5월 AI 콘퍼런스 소개



O'Reilly Artificial Intelligence Conference

비즈니스를 위한 AI의 혁신 및 모범 사례를 공유하는 콘퍼런스입니다. 주요 기업들의 AI 활용 사례와 구현 프로젝트, AI가 비즈니스와 사회에 미치는 영향에 대해서 다룹니다. 이론과 실제 산업 적용의 격차를 줄이는 것에 중점을 두고 있습니다.

날짜 : 4/29 ~ 5/2

장소 : 미국 뉴욕

홈페이지 : <https://conferences.oreilly.com/artificial-intelligence>

ICLR(International Conference on Learning Representations)

머신러닝의 성능은 데이터 표현의 선택에 크게 영향을 받습니다. 구성 모델링, 구조적 예측, 강화 학습 및 대규모 학습, 최적화 등 광범위한 주제들을 다룹니다. 이 기술들은 컴퓨터 비전, 음성인식, 텍스트 이해, 게임, 음악 등 다양한 분야에 적용되며, 특징을 학습하고 표현하는 방법들에 대한 연구 결과를 소개합니다.

날짜 : 4/30 ~ 5/3

장소 : 캐나다 밴쿠버

홈페이지 : <https://iclr.cc/>

ODSC(Open Data Science Conference) East

가장 큰 응용 데이터 사이언스 콘퍼런스 중의 하나로, 오픈 소스 툴과 라이브러리 및 언어 분야의 핵심 전문가들이 연사로 참여합니다. 해당 콘퍼런스는 산업에서의 데이터 과학의 실질적인 적용에 중점을 두고 있으며, 다양한 산업 분야에서 데이터 과학자, 소프트웨어 엔지니어, 분석가 등의 전문가들이 참여하는 행사입니다.

날짜 : 5/1 ~ 5/4

장소 : 미국 보스턴

홈페이지 : <https://odsc.com/boston/schedule>

AI CONGRESS LAS VEGAS

미국 포춘(Fortune)지 선정 500대 기업과 관련 회사들을 통해 AI와 머신러닝이 사업 방식을 변화하는 방법들에 대해 다룹니다. AI가 효과적으로 수익을 창출하는 방법과, 비용 절감 사례, AI 적용의 리스크와 도전 과제 등 주요 분야 전반에 혁명을 일으키고 있는 내용들을 설명할 예정입니다.

날짜 : 5/3 ~ 5/4

장소 : 미국 라스베이거스

홈페이지 : <https://theaicongressvegas.com/>

Train AI

실제 비즈니스 문제에 AI를 적용하는 방법을 배울 수 있는 행사입니다. 실제 현장에서 AI 적용 경험과 사례 연구, 솔루션 데모 등이 진행됩니다. 딥러닝 기술에 대한 교육과 실무자 대상의 실습, 회사 주요 경영진들의 워크숍이 함께 진행됩니다.

날짜 : 5/9 ~ 5/10

장소 : 미국 샌프란시스코

홈페이지 : <https://www.figure-eight.com/train-ai/agenda/>

Strata Data Conference

전 세계 데이터 과학자 및 분석가, 주요 경영진을 대상으로 전문가들의 데이터 사례 연구와 입증 결과, 새로운 분석 접근법 및 효과에 대해서 공유하는 자리입니다. 머신러닝, 데이터 엔지니어링 및 아키텍처, 빅데이터, 사이버 보안, 법률 및 윤리 등 다양한 주제를 다룹니다.

날짜 : 5/21 ~ 5/24

장소 : 영국 런던

홈페이지 : <https://conferences.oreilly.com/strata>

Rework Deep Learning Summit

전 세계 전문가들이 딥러닝 및 스마트 AI를 다룹니다. 음성 및 이미지 인식, 빅데이터, 신경망(neural network) 등의 기술이 비즈니스와 사회에 미치는 영향력을 알아보고, 글로벌 전문가들에 의한 최신 기술 발전 및 업계 동향을 공유합니다.

날짜 : 5/24 ~ 5/25

장소 : 미국 보스턴

홈페이지 : <https://www.re-work.co/events/deep-learning-summit-boston-2018>

Machine Learning Innovation Summit

업계 리더들이 비즈니스를 어떻게 변화시키고, 머신러닝을 적용함으로써 혁신적인 신제품을 개발하는지 그 방법들에 대해 공유합니다. 그리고 각 기업의 사례 연구를 통해 비즈니스를 최적화하고 성과와 효율을 극대화하여 제품을 생산하는지 그 방법들에 대해 공유하는 자리입니다.

날짜 : 5/9 ~ 5/10

장소 : 미국 샌프란시스코

홈페이지 : <https://theinnovationenterprise.com/summits/machine-learning-innovation-summit-san-francisco-2018>

LDV Vision Summit

컴퓨터 비전, 머신러닝, AI 등 다양한 분야의 트렌드와 기술을 논의하는 행사입니다. 전 세계 기술 혁신가들을 통해 비즈니스 현장의 생생한 소식을 들려줄 예정입니다. 기술적 주제들을 심도있게 다루는 ‘Technology Deep Dive’ 일정과 산업 생태계 현황을 다루는 ‘Business and Products’ 일정으로 양일간 진행됩니다.

날짜 : 5/23 ~ 5/24

장소 : 미국 뉴욕

홈페이지 : <http://www.ldv.co/visionsummit>

Data Science Summit 2018

eBay 및 IBM 공동 주관으로 진행되는 행사입니다. 학계 및 업계의 연구원, 데이터 과학자 및 개발자들이 참여하여 최신 AI, 데이터 과학, 머신러닝 응용 등을 논의하며, AI의 미래와 혁신, 빅데이터 분석 및 헬스케어, IoT, 교육, 핀테크 등의 분석 사례들을 다루는 트랙으로 구성됩니다.

날짜 : 5/28

장소 : 이스라엘 텔아비브

홈페이지 : <https://events.bizzabo.com/DataScienceSummit2018/home>

마치며

올해는 지난해보다 유난히 쌀쌀했던 벚꽃 시즌을 맞이했습니다. 이제는 새로운 초록의 잎사귀들이 돌아나는 봄이 시작될 것으로 예상됩니다. 발간사에서 말씀드린 바와 같이, 편집진은 4,5월에 걸쳐 '일상'이란 테마 속 인공지능이 변화시키고 있는 다양한 양상에 관한 총 5개의 글을 소개할 계획입니다. 우선 4월호에서는 여가, 노동 그리고 소비와 관련된 내용을 소개해 드렸습니다. 다음 호에는 일상의 두번째 파트로 교육과 부동산에 대한 소재로 찾아 뵙도록 하겠습니다. 날씨와 온도 뿐 아니라, 독자 여러분들의 마음 속 까지도 따뜻한 봄으로 가득 차길 희망합니다.

카카오 시리포트 원고 공모 REPORT

원고 마감일 | 매월 25일(익월호에 게재, 내부 판단에 따라 일정 변동 가능)
공모작 혜택 | 편당 50만 원
원고 주제 | 기술 전반 그리고 AI 산업 및 시장 분석
분량 | A4 5매 안팎(글자 크기 11pt, 행간 160%).
심사 | 카카오 시리포트 편집진과 카카오 AI 담당 기술 인력이 진행
투고처 | kakaoreport@kakaocorp.com
저작권 | 저작권법에 따른 정당한 범위 내에서 카카오 시리포트에 채택된 원고를 이용·가공·인용하시는 경우에도 반드시 출처를 밝혀주시기 바랍니다.

