

인공지능(AI) 개요 및 기술 동향

- 딥러닝(Deep Learning) 기술의 발달을 중심으로 -

(보안연구부 보안기술연구팀 / 2016.8.26.)

□ 개요

- 기술 연구 및 투자의 장기간 침체가 있었던 인공지능(AI)은 최근 딥러닝 기반 기술의 발달 및 기존 기술과의 결합 등을 통해 산업 전반에 적용 가능한 수준으로 발전하고 있음
- 이에 현재 활용되고 있는 인공지능 기술 중 딥러닝에 대한 개념 및 기술 동향을 중심으로 소개하고자 함

□ 인공지능(Artificial Intelligence, AI)

- (개념) 인간의 지능(인지, 추론, 학습 등)을 컴퓨터나 시스템 등으로 만든 것 또는 만들 수 있는 방법론이나 실현 가능성 등을 연구하는 기술 또는 과학
 - ※ 초기 인공지능의 개념은 미국 다트머스(Dartmouth) 학술회의에서 존 매카시(John McCarthy)가 “기계를 인간 행동의 지식에서와 같이 행동하게 만드는 것” 이라고 제안(1956년)
- (구분) 인공지능의 목표에 따라 4가지, 실제 사고해결 유무에 따라 2가지로 구분
 - (인공지능의 목표에 따른 구분) 인공지능의 목표에 따라 아래 그림과 같이 분류

< 인공지능의 목표에 따른 분류 >

| | | |
|------------------------|---|--|
| 인간의 사고작용 (thinking) | 인간과 같은 사고 시스템 (1985년 Haugeland와 1978년 Bellman의 정의) | 합리적 사고 시스템 (1985년 Charniak과 McDermott, 1992년 Winston의 정의) |
| | 인간과 같은 행동 시스템 (1990년 Kurzweil과 1991년 Rich와 Knight의 정의) | 합리적 행동 시스템 (1990년 Schalkoff와 1993년 Lugar와 Stubblefield의 정의) |
| 행동(behavior) | 이론적(ideal) | 합리적(rational) |

자료 : 조영임, 인공지능 기술 동향 및 발전 방향, 정보통신기술진흥센터(2016.2.17) 자료 재구성

- (인공지능의 사고 해결 유무에 따른 구분) 최근 인공지능의 사고 해결 유무에 따라 약한 인공지능(Weak AI)과 강한 인공지능(Strong AI) 2가지로 구분

< 인공지능의 사고 해결 유무에 따른 구분 >

| | |
|---------|--|
| 약한 인공지능 | 어떤 문제를 실제로 사고하거나 해결할 수 없는 컴퓨터 기반의 인공적인 지능을 만들어 내는 것에 대한 연구이며, 학습을 통해 특정한 문제를 해결 |
| 강한 인공지능 | 실제로 사고하거나 해결할 수 있다는 점에서 약한 인공지능과 차이가 있으며, ①인간의 사고와 같이 컴퓨터 프로그램이 행동 및 사고하는 인간형 인공지능과 ②인간과 다른 형태의 사고능력을 발전시키는 컴퓨터 프로그램인 비인간형 인공지능으로 구분 |

□ 인공지능 기술

- **(재부상 배경)** 최근 ①빅데이터 발달, ②정보처리(연산, 저장) 능력의 향상, ③딥러닝 알고리즘의 향상, ④클라우드 기반 환경 등의 영향으로 인공지능의 학습, 추론, 인지 기술을 발달시킬 수 있는 환경이 조성됨에 따라 재조명되기 시작함
- **(주요 기술)** 전산신경과학(Computational Neuroscience), 로봇틱스 인지로봇 공학(Robotics) 등 다양한 학문 영역에 걸쳐 있고 정형화된 기술 분류 체계는 아직 미흡하지만, 최근 ‘패턴인식’, ‘머신러닝’, ‘딥러닝’ 등이 인공지능 관련 기술 중 주요 기술 분야¹⁾로 언급됨
- **(기술의 특성)** 인공지능은 일반성, 방대성 등 지식의 특성뿐만 아니라 일반 소프트웨어 시스템과 달리 추론 기능 등의 특성을 가짐

< 인공지능 기술의 특성 >

| 특성 | 내용 |
|----------|---|
| 일반성 | 중요 성질을 공통적으로 가지고 있는 상황을 그룹화 함 |
| 방대성 | 중요 성질이 공통적으로 형성되기 이전의 기초자료는 매우 방대함 |
| 부정확성 | 현실문제에서 주어지는 정보가 부정확하며, 그로 인해 표현 역시 불명확함 |
| 지식 이용 | 소프트웨어는 자료나 정보를 사용하는 반면 인공지능은 인간과 같은 지식을 이용 |
| 추론 기능 | 적은 자료로 해(解)를 찾거나 많은 자료에서 공통점을 찾아내는 것 등은 인간의 사고과정에서 발생하는 추론기능과 같음 |
| 휴리스틱 탐색 | 경험적인 방법의 원리를 이용하여, 방대한 양의 탐색영역에서 최적의 해를 구하고자 일정한 규칙에 근거하여 해를 찾는 방법인 휴리스틱 탐색(heuristic search) 기능 제공 |
| 출력효율성 제고 | 입력정보가 비교적 완전 혹은 완벽하지 않더라도 결과물을 생성하는 기능 |

자료 : 석왕헌, 이광희, ETRI, 인공지능 기술과 산업의 가능성, 2015.10.30.

□ 딥러닝(Deep Learning) 기반 인공지능 기술

- **(딥러닝²⁾의 혁신성)** 인공지능 기술 발달의 핵심은 학습, 추론 등을 수행하는 인간 뇌의 특성을 컴퓨팅 모델화하는 것이며, 인간의 신경망 구조를 모방한 딥러닝은 초고용량 학습 알고리즘의 특징을 통해 기존 알고리즘과 달리 고성능을 발휘

1) ‘패턴인식’, ‘머신러닝’, ‘딥러닝’에 대한 개념은 「전자금융과 금융보안」 창간호(2015.7)의 “머신러닝을 활용한 스마트 서비스와 금융” 참조

2) 상세한 ‘딥러닝의 개념 및 특징’은 금융보안원, “머신러닝 개요 및 활용 동향”(2016.3.24.) 보고서에서 참조

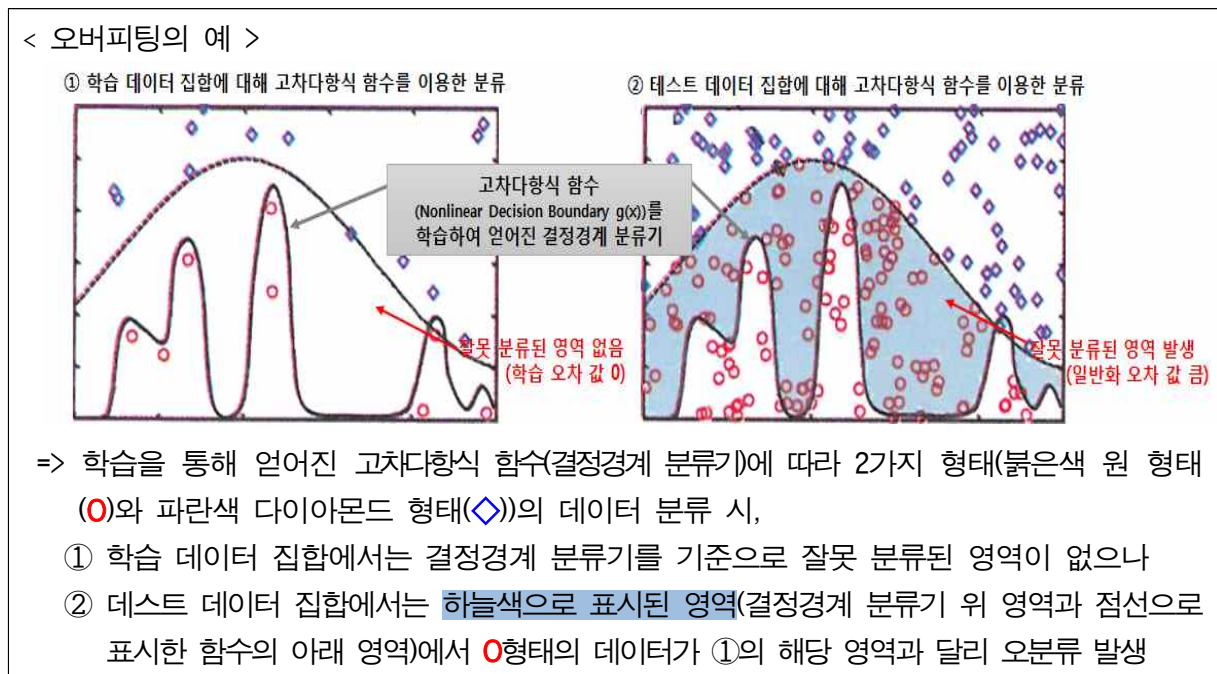
- (딥러닝의 특징) ①특징 추출(Feature Extraction)을 위한 전처리 과정(pre-processing)을 전체 학습 프로세스에 포함*, ②영상 데이터와 같이 차원수가 크고 복잡한 데이터의 경우 전처리 과정에서 손실될 수도 있는 정보를 자동 추출, ③높은 수준의 추상화 작업(Abstraction)**에 효율적이며 시물레이션 크기를 늘릴수록 대량의 데이터를 흡수하는 능력이 좋아지는 특징

* 기존의 전처리 과정(주어진 데이터를 다루기 쉬운 형태로 변환하는 과정)과 특징 추출(데이터에서 가장 핵심적인 정보를 특징으로 찾는 과정)의 두 단계를 분리하여 문제를 해결하던 방식을 하나의 프로세스로 통합

** 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 작업

- (딥러닝에서 오버피팅(overfitting)*의 최소화) 인공지능 등 학습에 있어서 나타나는 주요 문제점 중 하나인 오버피팅은 딥러닝 기술에서 강력한 형태로 방지하고 인공지능의 연구 개발을 가속화하는 중요한 계기를 마련

* 과(다)적합이라고도 하며, 만들어진 모델의 성능이 학습데이터(training set)에서는 좋지만, 새로운 데이터(test data)에서는 좋지 않은(혹은 일반화되지 않은) 경우를 의미함

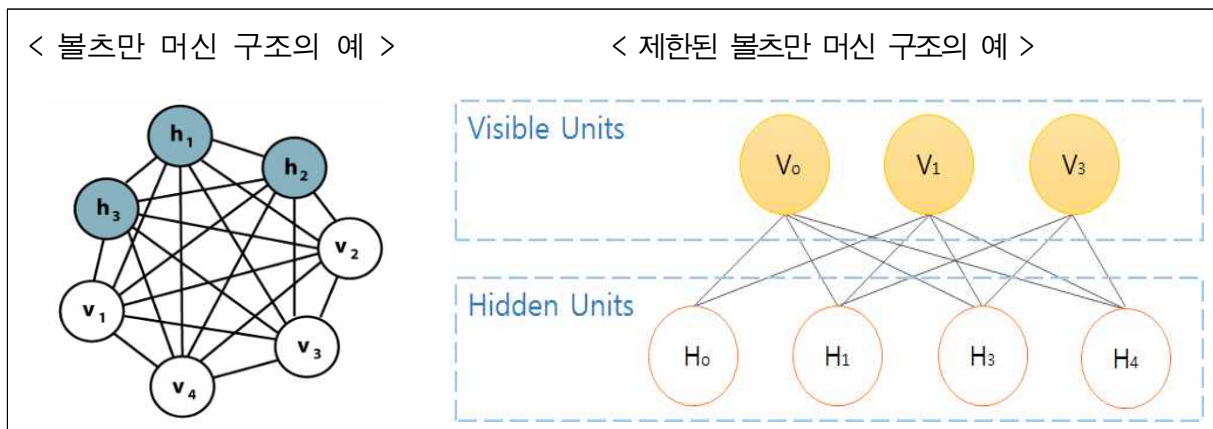


자료 : 박해영, 이관용, 패턴인식과 기계학습 재구성

- (오버피팅의 발생원인) 일반성 및 유연성 등 수많은 상황을 동시에 처리하는 인간의 뇌(지능)를 기존의 컴퓨터 구조로 구현하는 것은 매우 어려운 문제이며, 인공 지능에서 ‘학습’을 효율적으로 수행하기 위한 기술(알고리즘 등)들이 복잡도(complexity)가 너무 높은(혹은 낮은) 모델의 학습, 학습데이터와 현실 세계의 데이터와의 차이 등으로 인해 오버피팅의 문제가 발생

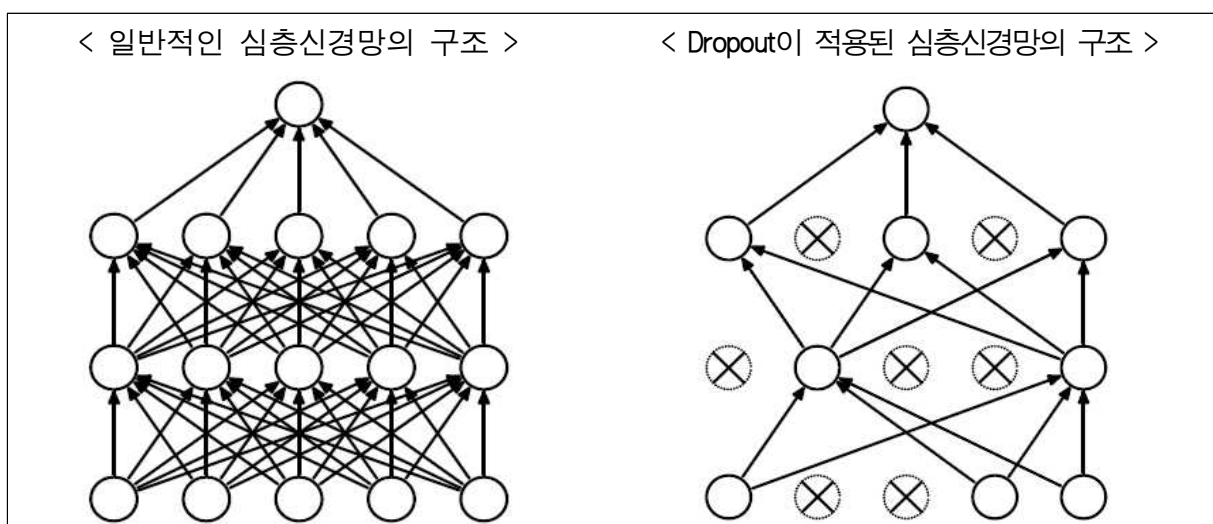
- (오버피팅 해결을 위한 딥러닝 알고리즘) 딥러닝은 ①제한된 볼츠만 머신 (Restricted Boltzman Machine, RBM), ②Dropout 형태 등의 **효율적 학습 형태**의 알고리즘을 통해 신경망의 일반화 능력 **최대화** 등으로 오버피팅을 최소화

- ① **제한된 볼츠만 머신**은 기존 볼츠만 머신에서 볼 수 있는 유닛(visible unit)과 은닉 유닛(hidden unit) 간 **연결망의 효율성을 고려하여** 제한을 둔 모델로써 두 개의 층을 구성하여 **신경망 학습의 효율성을 높임**



- ② Drop-Out은 학습단계에서 심층신경망(Deep Neural Network)의 모든 뉴런을 사용하지 않고 랜덤하게 보통 50% 정도만 사용하며 여러 개 신경망의 앙상블 효과*를 통해 오버피팅을 줄일 수 있는 방법임

* 학습 알고리즘들을 따로 사용하지 않고 다수의 학습 알고리즘을 유연한 구조로 만들어 더 좋은 성능을 얻는 효과



자료 : Nitish Srivastava, Georey Hinton et al., Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research 15(2014), p.1930, 2014.6.14.

□ 인공지능 기술의 동향 및 전망

- 글로벌 IT기업들은 인공지능 기술 및 플랫폼 기술을 육성하고 시장 영향력을 보다 확대하기 위해 주로 딥러닝 기술 위주의 인공지능 기술을 오픈 소스로 공개하는 추세

< 글로벌 IT 기업의 인공지능 기술 오픈소스 공개 사례 >

| 기업(단체)명 | 인공지능 기술 | 내용 |
|----------------------------|--|---|
| 구글 (Google) | 텐서플로우 (TensorFlow) | <ul style="list-style-type: none"> 구글 제품에 이용되고 데이터 플로우 그래픽 방식(Data Flow Graph)*을 사용하는 머신러닝(딥러닝) 오픈소스 라이브러리 일부를 공개(' 15.11.9) * 수학적계산과 데이터 흐름의 노드(Node)와 엣지(Edge)를 사용한 방향 그래프로 표현 |
| 페이스북 (Facebook) | 토치(Torch) | <ul style="list-style-type: none"> 오픈소스 기반 딥러닝 SW 프로젝트로 일부 딥러닝 모듈을 공개(' 15.1.16) |
| | 빅서(Big Sur) | <ul style="list-style-type: none"> 데이터를 학습 시 사용되는 인공지능 서버이며 오픈소스로 공개(' 15.12.10.) 엔비디아(NVIDIA) GPU가 8개가 장착되어 데이터 처리속도가 뛰어남 |
| 마이크로 소프트 (Microsoft) | CNTK | <ul style="list-style-type: none"> 코타나(Cortana, 인공지능 도우미)나 Skype 번역 어플리케이션에 사용하는 딥러닝 프레임워크인 Computation Network Toolkit를 오픈소스로 공개(' 15.4) |
| | DMTK | <ul style="list-style-type: none"> 분산 머신러닝 툴킷인 Distributed Machine Learning Toolkit을 오픈소스로 Github에 공개(' 15.11.12.) |
| | 프로젝트 AIX | <ul style="list-style-type: none"> MS의 인기게임인 '마인크래프트(Minecraft)'를 활용하여 인공지능 알고리즘을 테스트*하는 소프트웨어를 GitHub에 오픈소스로 공개(' 16.7.8.) * 실제 기계장치를 만들지 않고도 물리적인 대상을 다룰 수 있도록 개발자들이 로봇에 내리는 명령과 수행 결과를 테스트 할 수 있도록 돕는 AI 플랫폼 |
| IBM | 시스템 ML (System ML) | <ul style="list-style-type: none"> IBM 머신러닝 기술인 '시스템 ML'을 '아파치 스파크(Apache Spark)'* 오픈소스 생태계에 공개 계획을 밝힘(' 15.7.6.) * IBM의 향후 10년간 중요한 오픈소스 프로젝트이자 빅데이터 분석 기술 |
| 야후 (Yahoo) | 카페온스파크 (CaffeOnSpark) | <ul style="list-style-type: none"> 야후의 플리커(Flickr, 온라인 이미지 공유 서비스)에 적용한 딥러닝 소프트웨어를 오픈소스로 공개(' 16.2.16.) |
| 바이두 (Baidu) | Warp-CTC | <ul style="list-style-type: none"> 바이두 실리콘 밸리 연구소에서 개발한 AI SW로 컴퓨터가 사람의 말을 인식하기 위해 필요한 머신러닝 기술에 적용되며 오픈소스로 공개(' 16.1.14.) 바이두의 딥 스피치(Deep Speech, 언어 인식 프로그램) 2에 사용됨 |
| | 패들패들 (PArallel Distributed Deep LEarning) | <ul style="list-style-type: none"> 분산환경을 위한 딥러닝 플랫폼으로 데이터 분석 및 모델링을 지원하며 소스코드와 관련문서가 깃허브에 공개 예정(' 16.9.30.) |
| 삼성전자 | 베레스(Veles) | <ul style="list-style-type: none"> 딥러닝 응용 프로그램 개발을 위한 분산형 플랫폼을 오픈소스로 |

| 기업(단체)명 | 인공지능 기술 | 내용 |
|--------------------|---------|---|
| | | 공개(' 15.11) • 개방형 범용 병렬 컴퓨팅 프레임워크인 오픈 CL 또는 GPU에서 수행하는 알고리즘을 C프로그래밍 언어를 비롯한 산업표준 언어를 사용해 작성할 수 있도록 하는 GPGPU 기술인 쿠다(CUDA)를 사용 |
| 오픈 AI (Open AI) | 범용 인공지능 | • 테슬라 등의 기업이 약 10억달러(약 1조2000억원) 규모의 연구 자금을 모아 비영리연구단체 설립(' 15.12.14.) • 전 세계 인류에 도움을 주는 범용 인공지능을 만드는 것이 연구단체의 목표임 |

자료 : 각 기업의 인공지능 기술 공개 사이트 및 관련 기사

- 인공지능 기술 분야 중 딥러닝 관련 기술이 가장 빠르게 확산될 것으로 보이며, 이미지 인식 기술이 다음으로 높은 시장을 형성할 것으로 전망
- 다만 딥러닝과 이미지 인식 기술의 매출 규모는 많은 차이가 있으며 2020년까지는 대부분 딥러닝 기술이 시장을 주도할 것으로 전망됨

< 인공지능 기술별 관련 매출액 전망 >

단위 : 백만달러

| 기술 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 연평균 성장률 |
|-----------------|-------|-------|-------|---------|---------|---------|------------|
| 인지컴퓨팅 | 10.9 | 11.6 | 12.4 | 13.2 | 14.2 | 15.0 | 6.5% |
| 머신러닝 | 13.6 | 14.8 | 16.2 | 17.9 | 19.8 | 21.4 | 9.6% |
| 딥러닝 | 108.5 | 221.1 | 453.1 | 929.5 | 1,904.6 | 3,884.9 | 104.5% |
| 응용프로그램 인터페이스 예측 | 20.3 | 22.9 | 26.4 | 30.8 | 36.0 | 43.4 | 16.4% |
| 자연어 처리 | 7.8 | 10.1 | 13.0 | 16.6 | 21.1 | 26.3 | 27.5% |
| 이미지 인식 | 24.4 | 32.6 | 44.7 | 61.7 | 85.7 | 119.1 | 37.3% |
| 음성 인식 | 14.6 | 16.6 | 19.3 | 22.6 | 26.4 | 30.9 | 16.1% |
| 기타 | 2.3 | 2.4 | 2.6 | 2.8 | 3.1 | 3.5 | 9.1% |
| 합계 | 202.5 | 332.1 | 587.5 | 1,095.1 | 2,111.0 | 4,144.7 | 82.9% |

자료 : 석왕현, 이광희, 인공지능 기술과 산업의 가능성, ETRI Issue Report 2015-04, 2015.10.30.의 Tractica 2015 자료 재인용

- 현재 딥러닝 기술의 발달을 기반으로 산업에 적용 및 기술 발전 중인 인공지능은 ‘약한 인공지능’에서 ‘강한 인공지능’ 형태로 진화할 것으로 예상됨
- OVUM³⁾에서는 딥러닝 기술의 발달과 더불어 인공지능 기술은 과거 5년간 제한된 범위에서 특정한 내용(예: 구글의 보이스 등)을 처리하기 위해 활용되었으나,

3) www.ovum.com, IT 및 통신에 대한 연구(Research) 및 컨설팅 회사

- 향후 10년간은 인공지능 시스템이 확장된 범위(예:지능형 홈시스템 등)에서 산업 시장에 활용될 것으로 예측

< 2025년까지 인공지능 기술이 적용될 신흥 및 개발 산업 시장에 대한 예측 >

| 구분 | | 2015-2017 | 2018-2021 | 2022-2025 |
|----------------|------------|---|---|---|
| 신흥 산업 시장 | 비즈니스 측면 | <ul style="list-style-type: none"> • 상업적 서비스로 IBM의 Watson과 같은 AI기반 의사결정 지원 시스템 • 의학/의료, 과학연구, 교육 등을 포함한 서비스 | <ul style="list-style-type: none"> • 로봇의 두뇌에 AI적용 • 사무실, 공장, 외부 등에서 AI가 적용된 로봇의 상용화 | <ul style="list-style-type: none"> • 딥러닝 알고리즘의 지속적인 개선 및 사용 증가 • 예술가가 새로운 가능성을 탐구하는 곳에서 획기적인 AI의 출현 |
| | 소비자 측면 | <ul style="list-style-type: none"> • AI기반 소비자 제품/앱(음악 인식 앱 등) • 예) 이미 상용화된 애플의 시리, 구글의 보이스 등 | <ul style="list-style-type: none"> • AI가 지원 및 자동화 시킨 지능형 홈 시스템 • 예) TV프로그램 녹화, 요리 및 게임할 수 있는 AI기반 홈 시스템 | <ul style="list-style-type: none"> • 시끄러운 환경에서도 음성 인식률이 99% 이상으로 정확도의 발달 |
| 개발 산업 시장 | 비즈니스 측면 | <ul style="list-style-type: none"> • 금융 알고리즘 기반 트레이딩 시스템 등 AI기반 시스템의 지속적 성장 | <ul style="list-style-type: none"> • 제품과 소프트웨어 앱의 표준 기능에 진보된 AI가 지원하는 시스템이 적용 | <ul style="list-style-type: none"> • 사무실과 가정의 로봇에 발전된 AI 적용의 일반화 |
| | 소비자 측면 | <ul style="list-style-type: none"> • 기존 게임에 포함된 AI기능 외에 독창적인 AI기능의 발달 | <ul style="list-style-type: none"> • 스마트 기기에서 사용하기 위한 지능형 개인 앱의 급격한 소비시장 형성 | <ul style="list-style-type: none"> • 지능형 홈 시스템은 AI시스템, 응용프로그램, 유지보수 요구사항 등에 대한 시장 기회 창출 |

자료 : OVUM, Machine Learning in Business Use Cases, 2015.4.20.

- 또한 한국전자통신연구원에서 발표한 바에 따르면 빅데이터와 딥러닝 기술 기반의 약한 인공지능('15년)에서 인간의 능력을 증강시키는 기술('18년)로 진화하며
- 스스로 사고·판단·예측, 스스로 학습·진화, 두뇌를 모사하는 인지컴퓨팅 등 강한 인공지능('25년) 기술로 진화될 것으로 전망⁴⁾

□ 시사점

- (인공지능 기술 중 딥러닝 기술과 융합한 응용 애플리케이션 개발 및 연구에 대한 분석 필요) 현재 산업에 적용 및 기술 발전 중인 딥러닝 기술을 중심으로 타기술과 융합 등을 통해 활용 가치를 분석하는 것이 선행적으로 필요
- (인공지능 플랫폼 활용) 인공지능을 차세대 플랫폼 기술로 육성하기 위한 글로벌 IT기업들의 전략 중 하나인 인공지능 플랫폼을 활용하여 금융 분야에 맞는 서비스 분석 및 개발을 통해 경쟁력 확보 필요

4) 김병운, 인공지능 동향분석과 국가차원 정책제언, 한국정보화진흥원 정보화정책 제23권 제1호(pp.74~93)의 자료 중 ETRI(2016) 재인용 자료