

# 인공지능 기술의 이해



# [인공지능 기술의 이해]

- ① 4차 산업혁명과 인공지능
- ② 인공지능 개론
- ③ 알고리즘 설명 및 활용

# 인공지능과 4차 산업혁명- 4차 산업혁명이란?

4차 산업혁명?

인공지능?

Academy는 왜?



# 산업혁명의 진화



산업혁명 : 기술의 개발로 산업 및 사회경제 전반적으로 일어난 획기적인 변혁.

- ex) 1차 산업혁명은 영국에서 18세기 초에서 19세기 중반에 기술 발명으로 인해 사회적 경제적 변화를 일으킴.



1차 산업혁명  
- 증기기관 발명  
(1836)



2차 산업혁명  
- 전기에너지 발명 및 대량 생산(1908)



3차 산업혁명  
- 컴퓨터 및 인터넷 기반 지식 산업  
(1969)

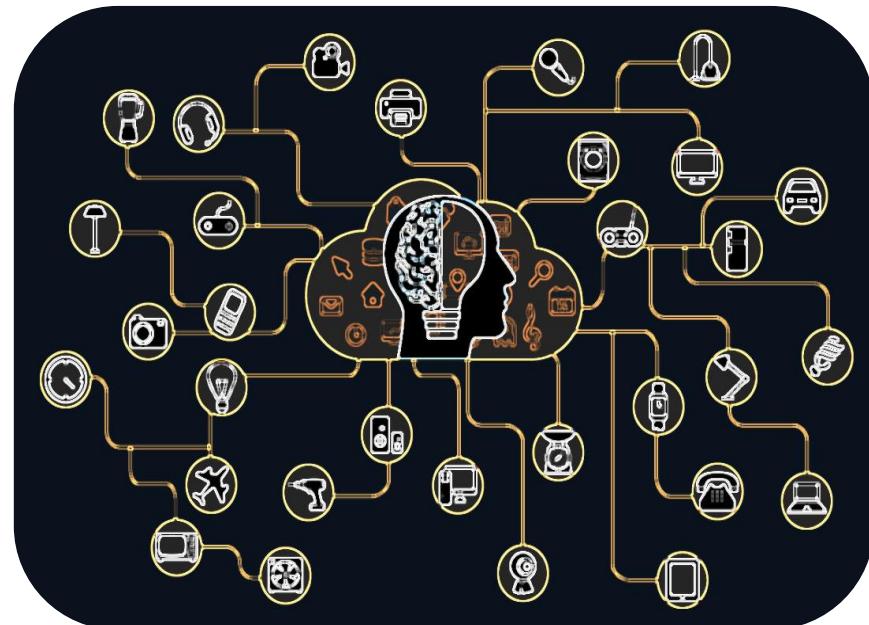
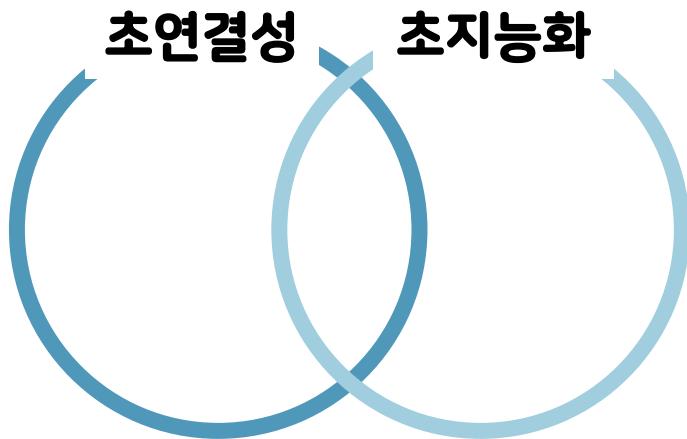


4차 산업혁명  
- 초연결성과  
초지능화의 결합  
(2016)

# 4차 산업혁명이란?



- 2016년 1월 다보스 포럼에서 제 4차 산업혁명을 정의
- 4차 산업혁명의 키워드 : CPS(Cyber-Physical System)  
‘물리적공간과 사이버 공간 간의 경계를 모호하게 하여서 융합시키는 시대’
- 4차 산업의 특징 : 초연결성과 초지능화의 결합



# 인공지능과 4차 산업혁명- 4차 산업혁명이란?

빅데이터  
클라우드 서비스

인공지능

제4차 산업혁명

3D 프린팅  
로봇산업

자율주행자동차  
스마트공장

바이오공학  
가상현실

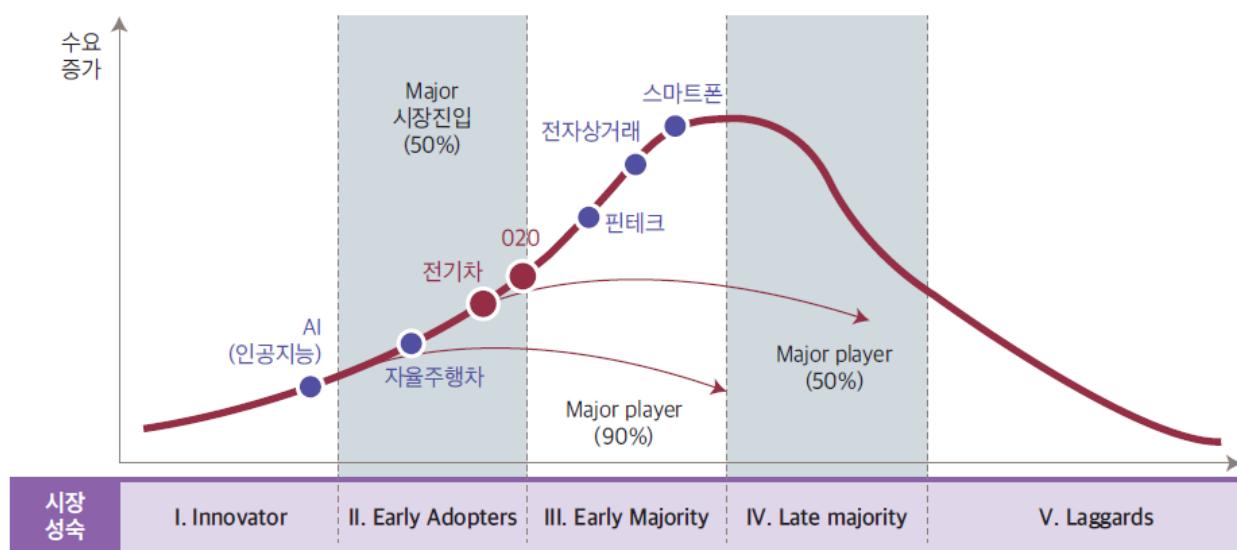


# 4차 산업혁명에서의 인공지능



- 인공지능 : 초연결성과 초지능성이라는 4차 산업혁명의 핵심을 모두 지니고 있음.
- 제품과의 결합을 통해서 다른 4차 산업을 확장 시킴
- 인공지능은 현재 도입기, 발전 가능성 농후

## 스마트산업내 분야별 라이프사이클



※ 출처 : Global Market Strategy(삼성증권, 2016)



## 가트너 선정, 미래 3개 기술 트렌드(Hype Cycle), 2019

### AI Everywhere

- 3~5년 이내에 최정점에 도달할 기술들은 대부분 인공지능 관련 분야.
- 여러 인공지능 탑재 기술 중에서도 자율주행 자동차로 점화
- Deep-learning, Deep-reinforcement learning, Artificial General intelligence, Autonomous Vehicles, Cognitive Computing, Commercial Drones, Conversational User Interface, Smart Robots, Smart Workspace, etc

### Transparently Immersive Experiences (순수몰입경험)

- 순수몰입경험은 사람과 기업, 사물간에 투명성 도입이 되어 연결되는 것을 이야기함.
- 기술 발전에 따라 관계의 복잡성이 증가될 전망
- 4D printing, Human Augmentation, Augmented Reality, Nanotube Electronics, Brain-Computer, Interface Virtual Reality, Volumetric Displays, Connected Home

### Digital Platforms

- 기술 인프라 중심의 플랫폼에서 '생태계 중심' 플랫폼으로 전환.
- 현재는 신빙성 있는 케이스 부족으로 인하여, 의사결정에 모험이 필요.
- 5G, Neuromorphic Hardware, Digital Twin, Quantum Computing, Edge Computing, Serverless Paas, Blockchain, Software-Defined Security, IoT Platform

# 가트너 Hype Cycle (2019)



## 가트너 Hype Cycle, 2019

- 가트너에서 매년 신기술 관련 리포트를 발표.
- 3~5년 이내에 최정점에 도달할 기술들은 대부분 인공지능 관련 분야.
- 인공지능도 흐름, 현재는 전문 시스템 위주 인공지능.

Gartner Hype Cycle for Emerging Technologies, 2017



# 4차 산업혁명 관련 기술 구성요소의 정의 및 역할

## 자능정보기술의 구성요소

인공지능



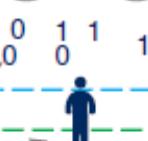
클라우드 컴퓨팅



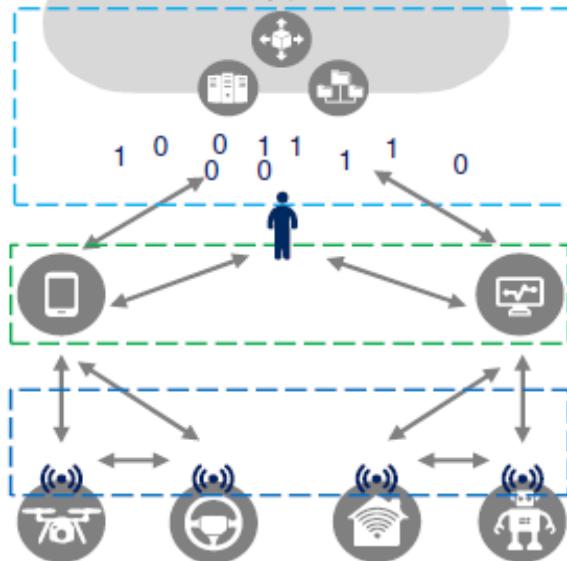
빅데이터



모바일



IoT  
(Internet of Things)



## 각 구성요소의 정의 및 자능정보기술에서의 역할

### 인공지능

- 인간의 지능을 구성하는 감각, 인지, 감성 능력을 컴퓨터 기술을 이용하여 활용하는 기술
- 빠르고 정확한 연산 능력을 통해 **최소한의 인간의 Input 또는 인간의 Input 전혀 없이 의사결정 실행이 가능함**

### IoT

- 기기에 부착된 센서를 통해 정보를 수집하고, 무선 네트워크를 통해 타 기기 및 컴퓨터와 연결하는 전체 플랫폼
- 센서 및 네트워크를 통해 객체간 정보의 **Input 및 Output의 최종 디바이스**로 활용됨

### 클라우드 컴퓨팅

- 서버, 저장 공간, SW 등 ICT 자원을 필요 시 인터넷을 통해 서비스 형태로 이용할 수 있도록 구동하는 기술
- 분산 저장 및 분산 연산을 통해 대용량의 **정보 처리가 신속하게 가능함**

### 빅데이터

- 이전에는 활용이 불가능한 규모의 데이터 모음 및 이를 수집, 저장, 관리, 분석하는 기술
- 대량의 정보가 실시간으로 생성되는 자능정보기술에서 **분석이 가능해지므로 정보의 효용이 증가함**

### 모바일

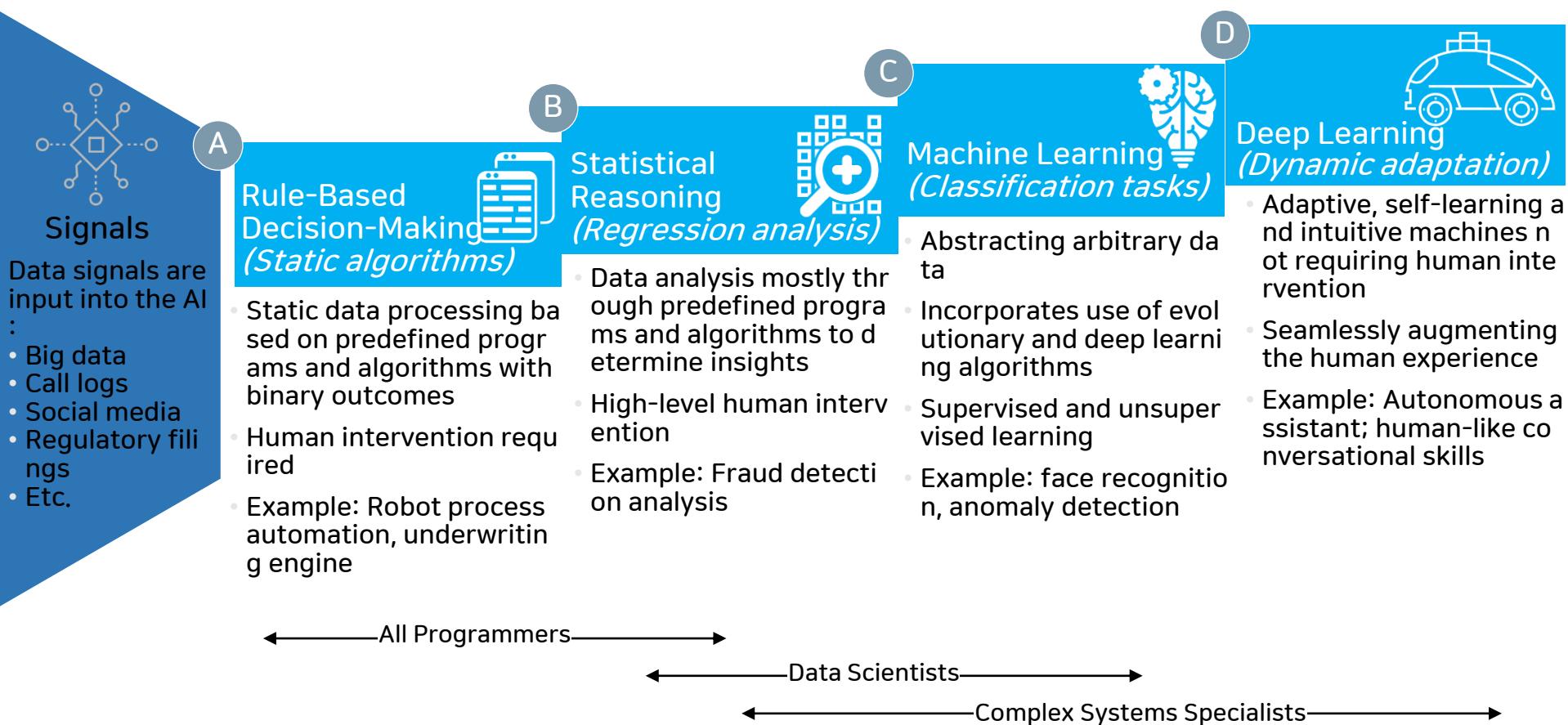
- 작업을 수행하거나 처리할 수 있는 이동형 하드웨어 및 정보의 교환을 위한 무선 네트워크 기술
- **정보 교류의 통로가 형성됨**

다섯 가지 구성요소의 상호작용을 통해 자능정보기술의 기반이 완성되어 가고 있음

# Stages in AI Advancement



최근 AI라는 용어는 Deep Learning 개념 중심으로 쓰이고 있으나, 실은 훨씬 광의의 개념임



# 인공 지능(Artificial intelligence)의 이해

빅데이터, 머신러닝을 지나 인공지능 기반의 디지털 혁신이 등장하고 있으며,  
이를 활용한 업무 혁신을 추진하기 위해 인공 지능 기반 기술에 대한 이해가 필요

## 인공 지능(Artificial intelligence) Framework

### Input

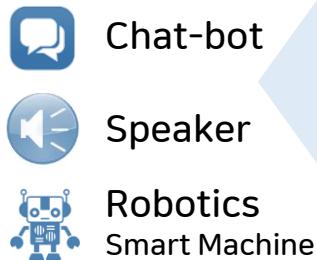


### **Perception Module**

언어 인지/대화 처리

- Data Extractor(STT/OCR)
- NLU
- Context Cognition
- Dialog Management

### Output



### **Action Module**

Legacy System과 연계되어  
실제 업무 처리 수행

- Legacy Application System
  - NLG
  - RPA
- Robotic Process Automation



### **Memory Module**

Knowledge Base 구축

- Rule Based
- Triple Wiki(색인/검색)
- Exo-brain
- MRC

### **Reasoning Module**

Advanced Analytics를  
활용한 의사결정 수행

- Inference Engine
- Predictive Model
- Decision Making

# 선진사 인공지능 활용 사례



글로벌 선진 보험사들은 다양한 영역에서 인공지능 등 디지털 新기술을 활용한 혁신활동을 추진하고 있음

## 인공지능 활용 가능 영역

## 활용 사례

## 세부 내용

단순 업무 지원 (Smart Advisor)	<ul style="list-style-type: none"><li>내부 직원을 대상으로 Help Desk 역할</li><li>AI가 최적의 답안을 추천해주지만, 사람이 최종적인 의사결정</li></ul>	<ul style="list-style-type: none"><li>대면 설계사 영업 지원</li><li>콜센터 상담사 지원</li><li>Claim 업무 지원</li></ul>	 	<ul style="list-style-type: none"><li>설계사와 고객의 대화 중 주요 키워드를 인식하여, 효과적인 대화를 위한 정보 제공</li><li>고객 문의 내용에 대한 응답 리스트와 필요 서류, 절차를 추천하여 상담 업무 지원</li><li>低경력 보상 담당자가 Claim 청구건 처리시, 조사 진행여부 판단을 지원</li></ul>
상담 업무 자동화 (Virtual Agent)	<ul style="list-style-type: none"><li>對고객 대상 24시간 상담서비스 제공</li><li>단순한 질문에 대한 FAQ 형태의 대화 서비스 제공</li></ul>	<ul style="list-style-type: none"><li>온라인 상담지원</li><li>콜센터 상담 효율화</li><li>보험 사고 접수 자동화</li></ul>	  	<ul style="list-style-type: none"><li>자동차 온라인 보험 가입 화면에서 보험 가입 단계의 대고객 상담 서비스를 제공</li><li>LINE 플랫폼을 활용하여 고객의 단순한 FAQ에 대한 자동 응대로 상담 업무 효율화</li><li>사고접수-처리-지급 업무를 고객이 전용 App을 통해 self처리 함으로서 비용 절감</li></ul>
업무 처리 자동화 (AI Automation)	<ul style="list-style-type: none"><li>AI 활용한 의사결정 자동화</li><li>전문적인 지식을 바탕으로 의사결정이 필요한 업무 자동화</li></ul>	<ul style="list-style-type: none"><li>상품 상담 및 보험 가입</li><li>U/W 인수 심사 자동화</li><li>Claim 보험금 지급 심사 자동화</li></ul>	  	<ul style="list-style-type: none"><li>보험 가입에 필요한 정보 항목에 대해 몇 가지 질문을 통해 여행자 보험 가입 처리</li><li>인수 심사 자동화를 통해 20분내 즉각적인 가입 처리 서비스 제공</li><li>Text Mining 기술을 활용하여 보험금 지급 업무를 자동화</li></ul>

# [인공지능 기술의 이해]

- ① 4차 산업혁명과 인공지능
- ② 인공지능 개론
- ③ 알고리즘 설명 및 활용

# 인공지능이란?



## 인공지능 정의

인공 : 사람의 힘으로 어떤 물체에 관해서 가공하거나 작용하는 일.

지능: 인간의 인식, 판단, 추론, 문제해결, 그 결과로서의 언어나 행동지령, 더 나아가서는 학습기능과 같은 인간의 두뇌 작용을 이해하는 것을 연구 대상으로 하는 학문 분야.

<참조: 위키백과>

## 인간의 두뇌 작용을 이해하고 **모방**하여 이를 기계에 적용하는 것

“의사결정, 문제해결과 같은 활동, 즉 인간의 사고와 관련된 활동의 자동화….”

(Bellman, 1978)

“인간이 더 잘하는 것을 어떻게 하면 컴퓨터가 하게 만들지를 연구하는 것”

(Rich and Knight, 1991)

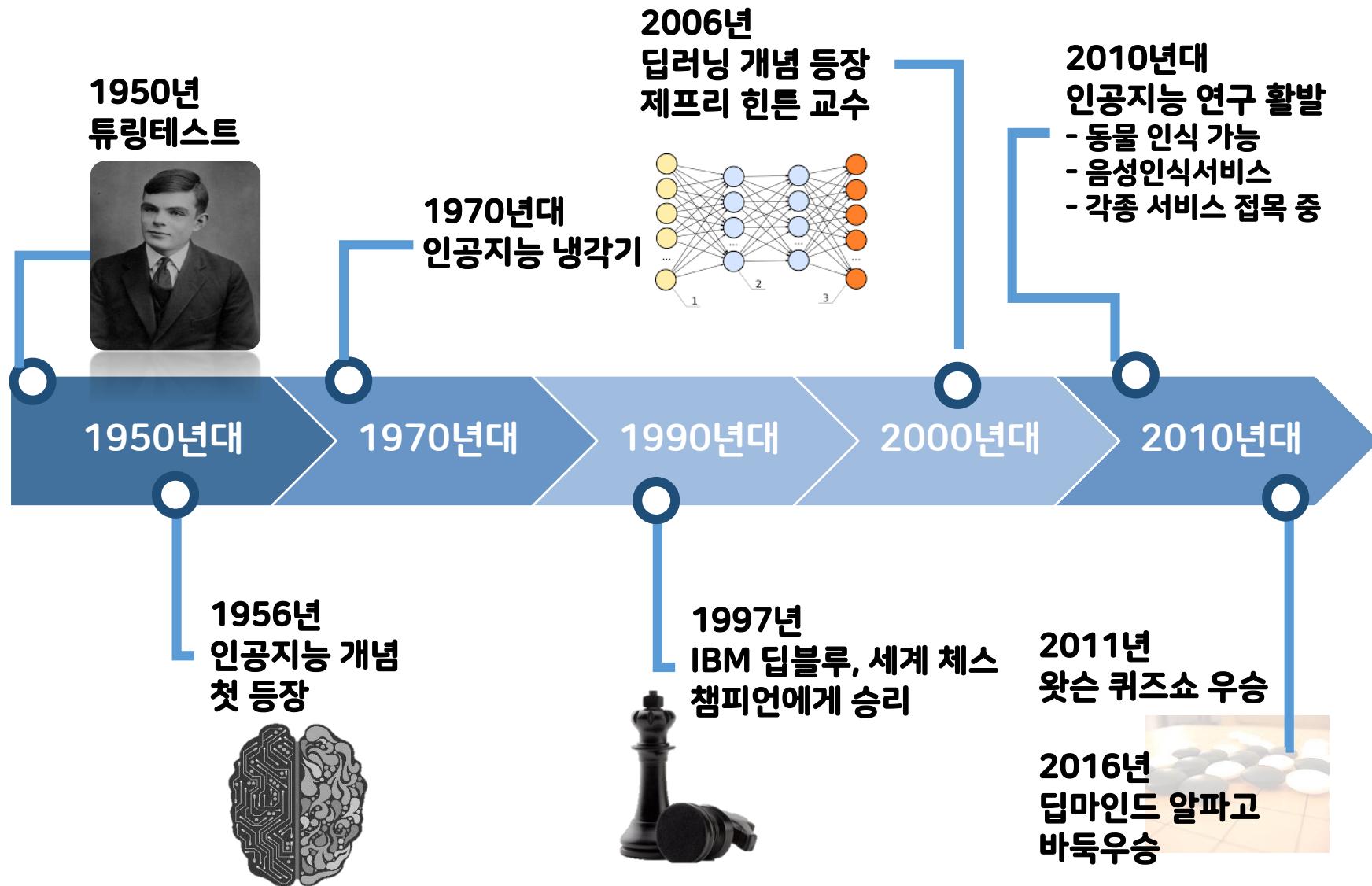
“계산적 모델의 사용을 통한 정신적 능력에 대한 연구”

(Charniak and McDermott, 1985)

“지능적 에이전트를 설계하는 것에 대한 연구”

(Pool et al., 1998)

# 인공지능의 역사- 인공지능의 발전 과정



# 인공지능의 종류- 강인공지능 vs 약인공지능

## 약인공지능(Weak AI)

- 약인공지능은 자아와 지성을 지니고 있지 않은 형태의 인공지능으로, 지능 중 극히 일부분의 기능들을 구현하고 있다.
- 규칙과 알고리즘 등의 정의된 과정에 따라서 결과값을 도출한다.

## 강인공지능(Strong AI)

- 강인공지능은 자아와 지성을 지니고 있는 형태의 인공지능으로, 자신이 스스로 판단하고 추론을 할 수 있다.
- 인간의 두뇌와 흡사한 기능을 하며, 한가지 기능에 얹매이지 않고 다수의 기능을 구현 할 수 있다.  
우리가 생각하는 인공지능의 모습과 비슷하다 할 수 있음.



<참조 : 지능형 콘텐츠 기술 발전 전략 연구 - 한국콘텐츠진흥원>

# 인공지능 3대 핵심 지능 및 분류



인공지능이란?: 인간의 두뇌 작용을 이해하고 모방하여 이를 기계에 적용하는 것



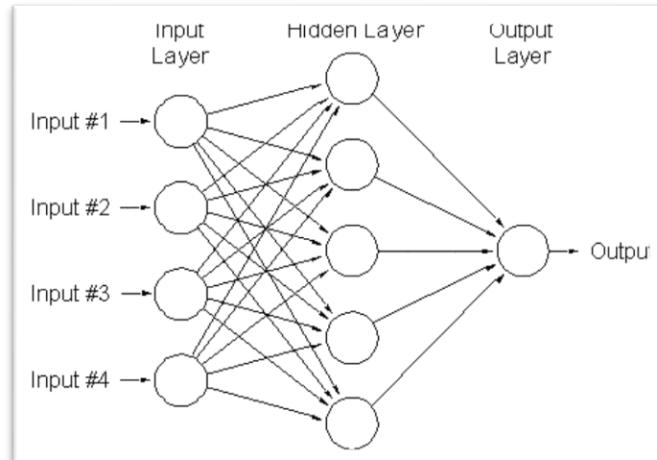


## 학습 능력(Learning ability)

- 학습 : 연습이나 경험의 결과를 통해서 일어나는 행동의 지속적인 변화 (출처 : 네이버사전)
- 인공지능에서의 학습능력이란, 정제된 데이터를 학습 알고리즘에 투입시켜 특징을 추출하도록 하는 과정.
- 다른 능력을 위한 가장 기본적인 능력.

Thinking Machines | Cognitive Computing | Deep Learning | Artificial Intelligence

A huge pr(e)omise:

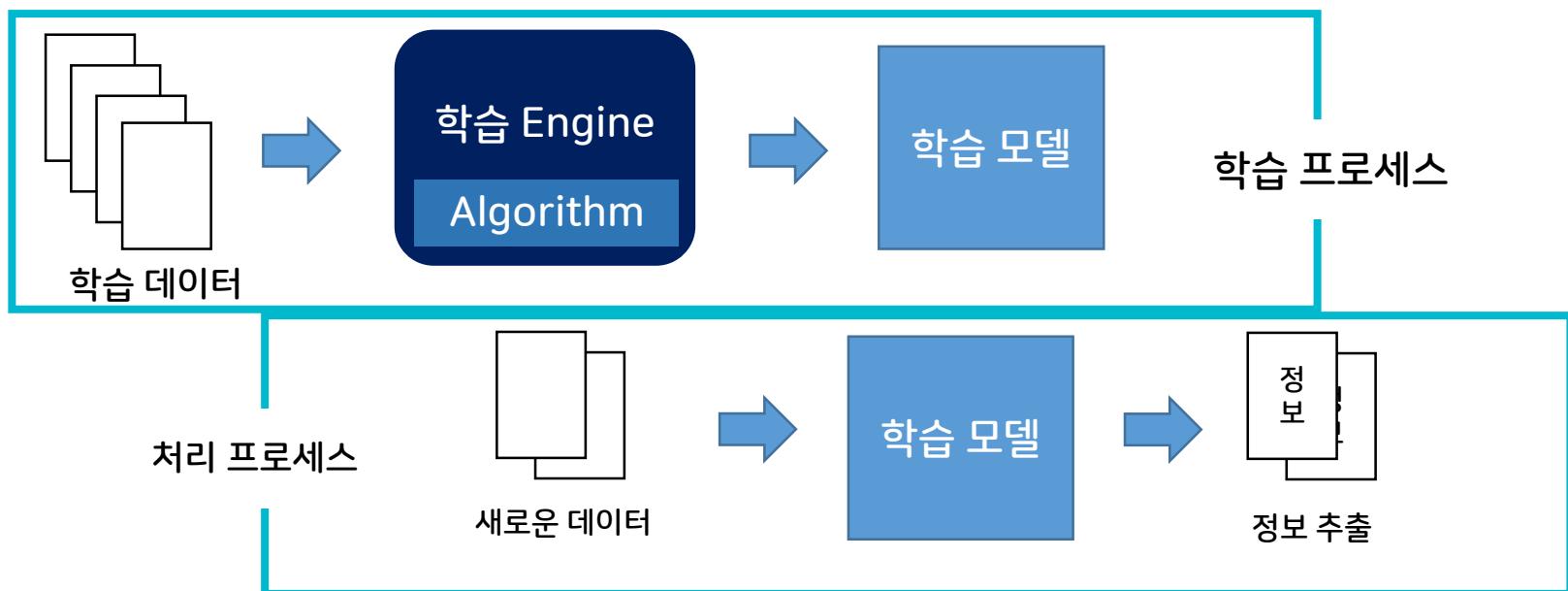


# 인공지능에서 학습의 의미란?



## 학습의 의미

- (정제된) 학습 데이터를 학습 알고리즘이 있는 학습 엔진에 넣어 학습 모델을 만드는 것.
- 학습엔진 역할 : 수많은 데이터 내에서 규칙성과 패턴을 찾아서 모델로 만들어 내는 것.
- 학습을 통해 만들어진 모델을 적용시켜서 새로운 데이터에 대한 정보를 추출.



# 학습의 능력을 좌우하는 것은 데이터, 그렇다면 데이터의 분류는?

## 정형 데이터(Structured Data)

- 고정된 필드에 저장된 데이터
- 수집되는 정보의 형태가 정하는 기준 : 데이터베이스를 설계한 기술자
- 한정된 분야에서 사용이 가능함  
Ex) 관계형 데이터베이스와 스프레드시트

## 비정형 데이터(Unstructured Data)

- 고정된 필드에 저장되지 않은 데이터
- 수많은 데이터의 약 85%가 비정형 데이터
- 실시간 정보들을 통해서 많고 다양한 기준으로 수집 가능  
→ 다양한 인사이트 도출  
ex) 페이스북, 트위터, 유튜브 영상, 이미지 파일, 음원파일 등

## 반정형 데이터(Semi-Structured Data)

- 고정된 필드에 저장된 데이터는 아닌, 메타데이터 및 스키마를 포함하는 데이터.
- 태그나 마커가 포함되어 시맨틱 요소를 구분 가능.
- 인터넷이 등장한 뒤로 증가되는 추세.  
ex) XML, HTML 텍스트등

# 인공지능 3대 핵심 기능 - 추론 능력



## 추론 능력(Reasoning ability)

- 추론 : 기존에 알려진 혹은 학습된 정보를 근거로 하여 새로운 정보 및 판단을 추출해 내는 행위
- 기존 데이터를 통해서 연관 관계를 파악하여 규칙을 발견.
- 대표적인 사례 :  
주가예측 인공지능, 고객이탈 분석, 고객 특성 분석, 추천 등



## 見性(Kensho)

컴퓨팅 "추론도 가능하다" AI 연구 새 이정표 써다

구글 딥마인드 연구팀, AI 관계형 추론 능력 증명

손경호 기자 | 입력 : 2017.06.09.15:58 | 수정 : 2017.06.09.16:42

1062

업계 최초 용량/성능/가용성보장 프로그램 제공! K2 GEN6 물플래시 스토리지 소개자료  
IBM 클라우드 매니저d 서비스 고객 사례 다운로드 – 타이탄플랫폼 브로서

지금까지 인공지능(AI) 기술의 한 분야인 딥러닝은 주어진 이미지나 텍스트를 인식하거나 문장을 다른 언어로 번역하는 등 작업에서 눈에 띄는 성과를 거두웠다.

그러나 여전히 인간처럼 여러 가지 상황을 종합적으로 판단해 다음 상황을 예측하는 수준의 지능을 갖는데 도달하지는 못했다.

지난 5일(현지시간) 구글 자회사 딥마인드 연구팀은 미국 코넬대 아카이브(arXiv.org)에 두 가지 논문을 새로 게재하면서 AI가 인간처럼 관계형 추론까지 할 수 있다는 사실을 증명했다.

연구팀은 '관계형 추론을 위한 단순한 인공신경망(A simple neural network module for relational reasoning)'이라는 논문에서 관계형 네트워크(Relation Networks, RNs)라는 개념을 선보였다.(논문링크)

연구팀은 "관계형 추론은 일반적으로 지능형 행동을 하기 위한 핵심 구성요소이지만 인공신경망을 학습시키는 방법으로 이를 입증하기는 어려웠다"며 "관계형 네트워크(Relation Networks, RNs)를 활용해 어떻게 이런 문제를 해결했는지를 설명한다"고 밝혔다.





## 인지 지능 (Cognitive Intelligence)

- 인지 : 사물을 분별하고 판단하여 아는 행위
- 인지 지능에는 언어, 시각, 청각, 동작이 존재
- 가장 흔한 인지 지능은 텍스트 인식과 음성 인식이 존재.
- 가장 기본이 되는 것은 패턴 인식.
- 인공지능에서 인지능력이란? 학습을 통해 발견된 패턴으로 현재 모습을 파악하는 것을 의미.

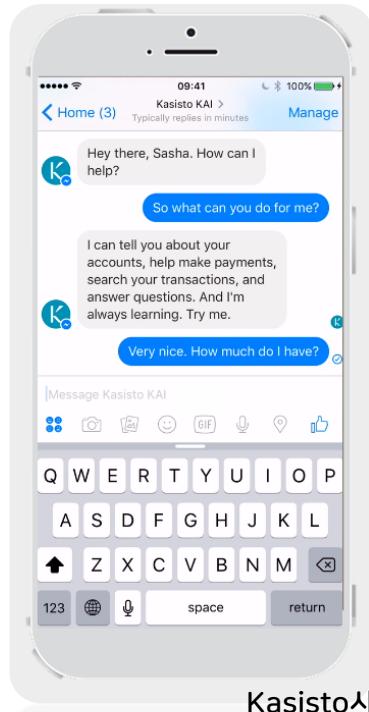


# 인공지능 3대 핵심 기능 - 인식 능력



## 언어지능(Linguistic intelligence)

- 컴퓨터가 텍스트를 인식하는 능력
- 컴퓨터가 언어를 이해 할 수 있게 하는 기술을 필요로 함.  
대표적인 기술 : 형태소분석, NER, Word-embedding 등
- 대표적인 사례 :**  
챗봇 사업, 인공지능 가상 상담원, 글자 인식 등



Kasisto사례



Maum.ai 사례



## 음성 인식능력 (Speech-recognition)

- 컴퓨터가 음성을 인식하는 능력
- 컴퓨터가 음성을 인식하고 해석 할 수 있게 하는 기술을 필요로 함.  
대표적인 기술 : STT, TTS 등
- 대표적인 사례 :  
인공지능 가상 상담원, 인공지능 스피커, 스마트 가전 등

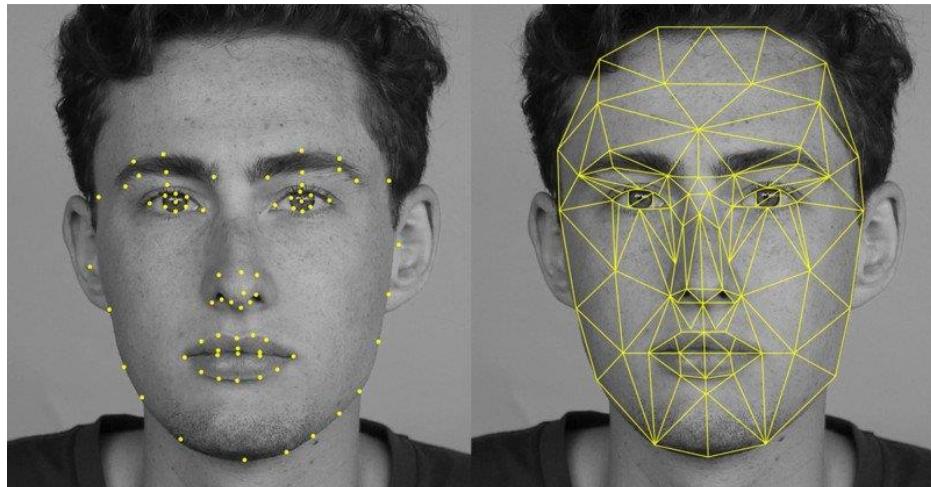


인공지능 스피커



## 시각지능(Visual Intelligence)

- 컴퓨터가 시각적으로 사물을 분별하고 인식하는 능력
- 컴퓨터가 학습을 통해서 형태에 대한 패턴을 발견하여 구분하게 됨.
- 선형맞춤과 비선형 변환을 반복해 쌓아 올린 구조.
- 대표적인 사례 :  
Amazon go, 구글렌즈, 자율주행자동차 등



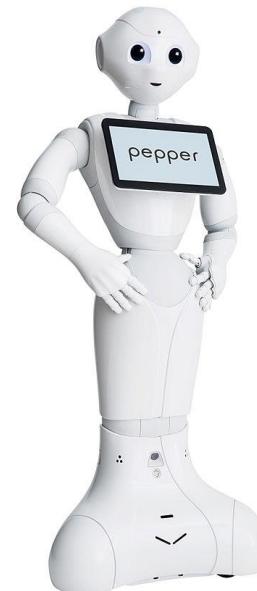
<출처 : petapixel>





## 동작지능(Robot Intelligence)

- 여러 센서기술을 통해서 대상자의 인식을 하여, 특정 행동을 제어하는 지능
- 센서기반 프로세스와 제어기반 프로세스로 나뉘어서 발전
- 센서는 인식과 컴퓨터 사이에서 중개자 역할을 수행
- 로봇공학과 인공지능과의 결합을 통한 시너지 효과.
- **대표적인 사례 :**  
Pepper, LEAP motion 등



# 인공지능의 장점과 한계



## | 인공지능의 장점과 한계

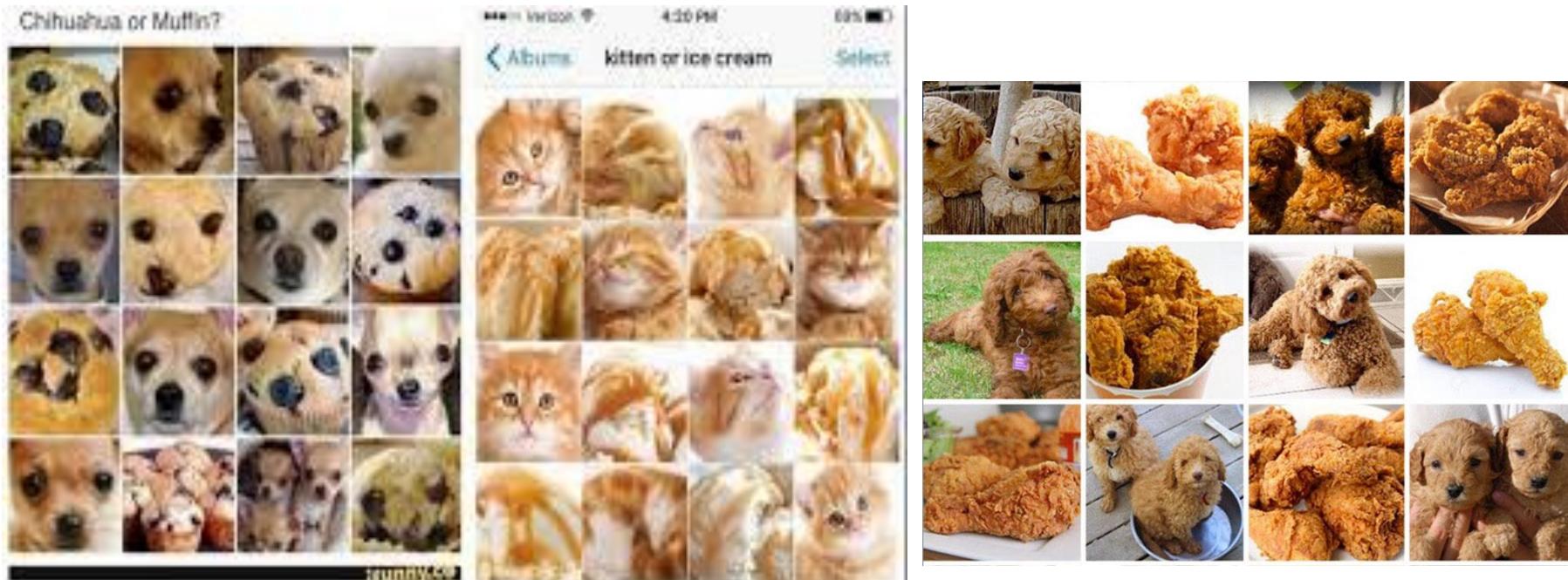
	인공지능 장점	인공지능 한계
내용	<ul style="list-style-type: none"><li>1) 인간의 개입 최소화</li><li>2) 개인능력 격차 완화</li><li>3) 비용절감</li><li>4) 자동화</li></ul>	<ul style="list-style-type: none"><li>1) 데이터 학습에 대한 한계.</li><li>2) 인공지능에 대한 법적 윤리적 책임소재 관한 사회적 규정 미비.</li><li>3) 판단의 기준 정의 필요.</li><li>4) 단일기능만 수행.</li></ul>

<참조 : Igloo security>

# 인공지능의 장점과 한계



## | 인공지능의 장점과 한계 - 학습의 한계



<출처 : google image>



## 자연어 처리(Natural Language Processing)

인간의 언어(텍스트)를 기계적으로 분석해서 컴퓨터가 이해할 수 있는 형태로 변환하는 일

- 문서를 문장으로 분할하고,
- 문장을 최소의 의미 단위인 형태소로 나누어 품사를 부여하고,
- 문장에 포함된 인물명, 기업명, 장소, 숫자 표현 등을 인식함



## 자연어 이해(Natural Language Understanding)

- 자연어 처리에서 한단계 진화된 형태, 단순히 자연어를 처리하는 것이 아닌 실제로 인간의 언어를 이해하기 위하여 처리하는 것.
- 형태소 분석, 개체명 인식 뿐만 아니라 구문 분석, 화자 의도 분석, 감성분석까지 결합시켜서 인간의 언어를 의도를 완벽히 이해하기 위함.

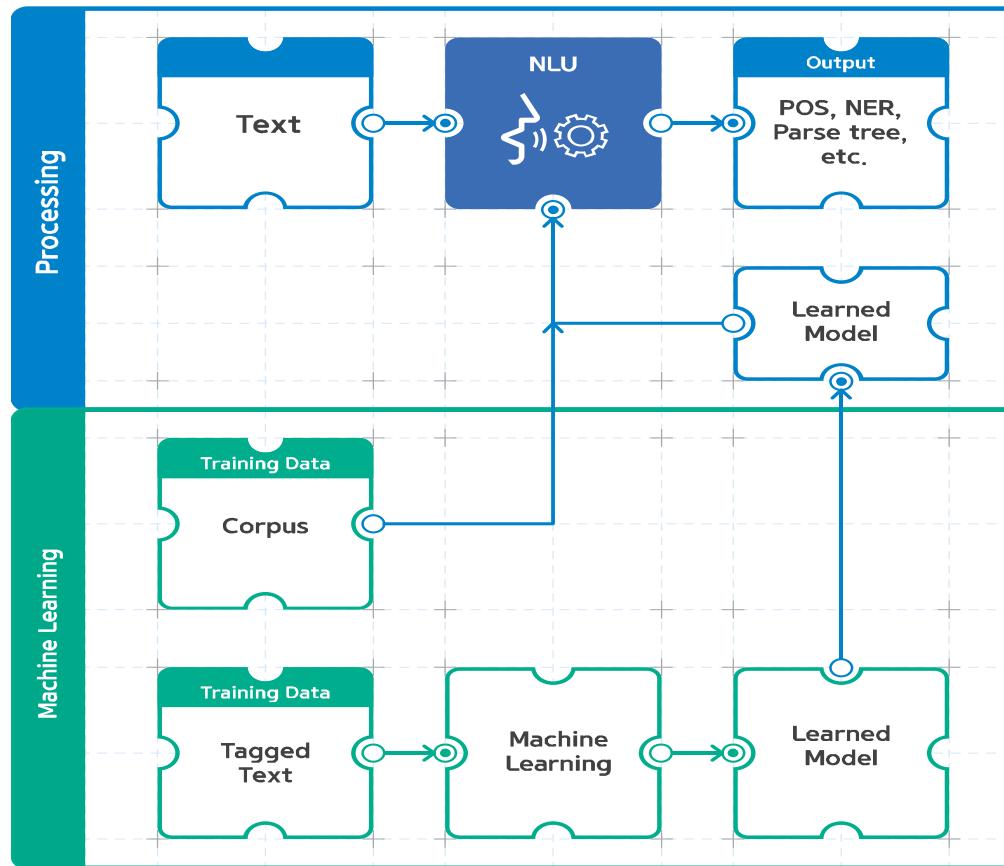


# 자연어 처리의 정의- 언어 이해 기술



## 자연어 이해(NLU)

- 미리 구축된 사전을 바탕으로 코퍼스에 태깅하는 형태로 학습데이터 형성
- 학습 데이터를 각 알고리즘으로 학습하여 학습 모델 형성



# 자연어 처리- 형태소 분석



## 형태소 분석(POS Tagging)

- POS는 Part-Of-Speech에 약자로, 형태소의 뜻과 문맥을 고려하여, 품사를 태깅해 주는 것.
- 형태소란? '뜻을 가진 가장 작은 말의 단위'
- 형태소 분석은 문장을 형태소 단위로 쪼개서 표현한 것을 이야기 함.
- 쪼갠 후 명사 및 동사 등의 품사를 태깅해 주는 역할까지 하게 됨.

ex) 미국 환율 알려줘

- 미국/미국/**NN**/조직\_지명 환율/환율/**NN**.용언불가능/양  
알리/알리/**VB**/+어주/어주/**EV**/+어/어라/**EE**.종결/명령형

순번	품사번호	품사명	품사명
1	0	nc	자립명사
2	1	np	대명사
3	2	nb	의존명사
4	3	nn	수사
5	4	pv	동사
6	6	pa	형용사
7	8	mag	일반부사
8	11	ii	감탄사
9	12	jc	격조사
10	13	co	지정사
11	14	ef	종결어미
12			DM



## 개체명 인식(Named-entity recognition)

- 단어의 정보에서 추출된 결과 값.
- 사람, 조직, 지명, 시간 등의 값을 미리 학습시켜 놓은 사전에 의해서 인식하여 추출하는 방법.
- 전처리와 후처리 방식 두 가지 존재.
- 전처리 사전과 규칙을 적용하여 개체 모호성을 사용자 주도로 해소.
- 모델 학습을 통하여 필터링 사전, 후처리 규칙과 후처리 사전을 적용하여 신조어 형태의 개체 유형도 효과적으로 탐지.

ex) 미국 환율 알려줘

- 미국/미국/NN/**조직\_지명** 환율/환율/NN.용언불가능/양  
알리/알리/VB/+어주/어주/EV/+어/어라/EE.종결/명령형

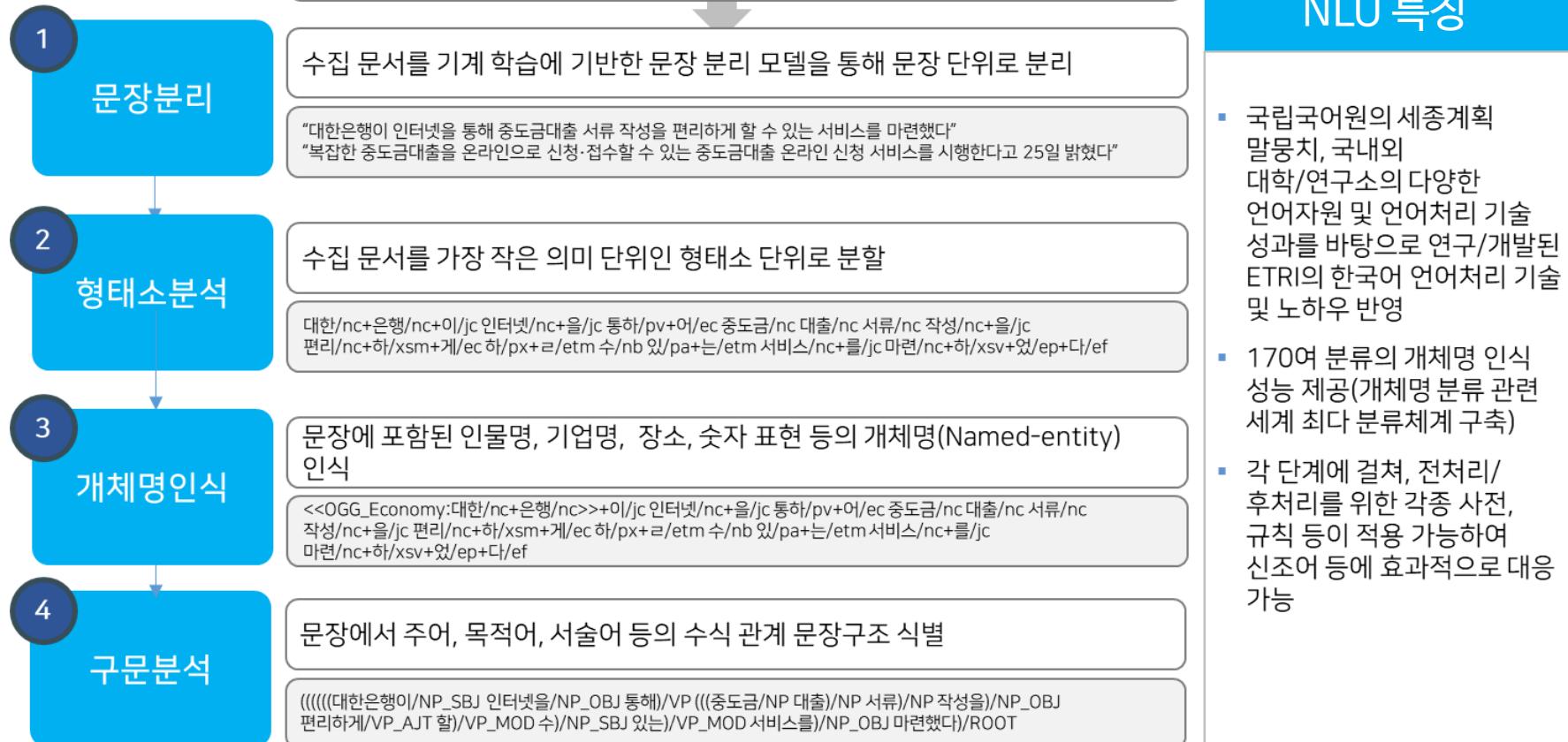
# 자연어 처리 기술 예시



## NLU Processing(Natural Language Understanding)

- 4단계로 구성.
- 문장분리/형태소분석/개체명인식/구문분석

“대한은행이 인터넷을 통해 중도금대출 서류 작성 편리하게 할 수 있는 서비스를 마련했다.  
복잡한 중도금대출을 온라인으로 신청·접수할 수 있는 중도금대출 온라인 신청 서비스를  
시행한다고 25일 밝혔다.”

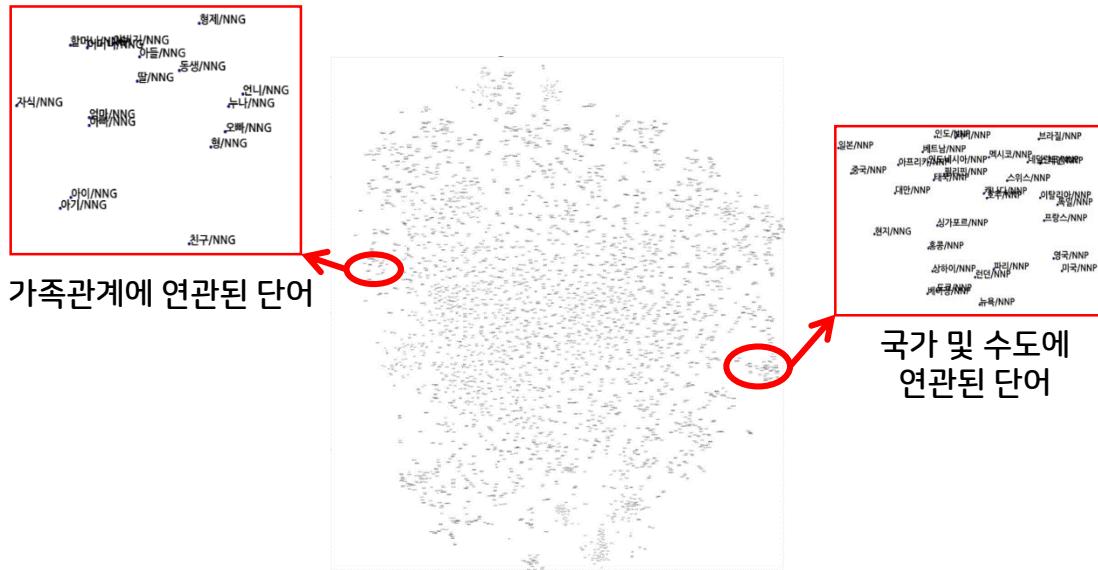




## 워드임베딩 기술(Word-Embedding)

### 워드임베딩(Word-Embedding)의 정의

- 대량의 문서 집합을 이용하여 단어 하나 하나를 수십 혹은 수 백차원의 벡터로 변환된 값
- 각 단어들 사이의 유사도를 측정 할 수 있고, 단어에 대해 수치적으로 쉽게 다룰 수 있음.
- 유사한 단어들끼리는 군집화(clustering)된다는 특징이 있음.



Word Embedding Visualization



## 워드임베딩 기술(Word-Embedding)

### 기존의 방식

- 컴퓨터가 어떤 단어를 인지할 수 있도록 수치적인 방법으로 단어를 표현
  - 단점 : 단어가 본질적으로 다른 단어와 어떤 차이점을 가지는지 이해 할 수 없음.
  - 단점 극복을 위해 단어 자체가 가지는 의미를 표현하는 방식으로 고안.
    - ✓ **One-hot Encoding**
      - 머신 러닝에서 활용할때 쓰이는 범주형 변수 지정 방법
      - 해당하는 칸의 정보를 1로 표시 나머지는 0으로 표시
- Ex) 경기도, 판교, 마인즈랩, 워드임베딩, 인공지능  
판교를 표현하기 위해서는?  
: [0, 1, 0, 0, 0]

# 워드 임베딩: 최근 방식



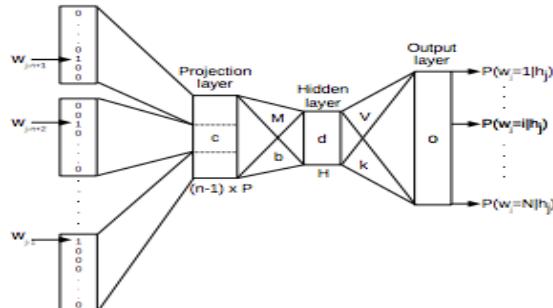
## 최근의 방식

- 최근에 들어와서는 Neural Network 방식을 이용한 워드임베딩 사용하기 시작함.
- 단어 자체에 의미를 부여할 수 있게 됨.
  - ✓ 대표적인 방법 : NNLM, RNNLM, Word2Vec, GloVe

### NNLM

#### (Feed-Forward Neural Network Language Model)

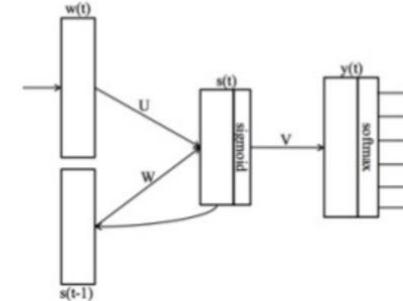
- NN기반 단어 학습 모델의 초기 컨셉
- 단어들에 대해서 one-hot encoding 벡터화.
- 특징 : 파라미터 고정, 이전 단어에만 국한 되어있다는 점, 연산이 느림.



### RNNLM

#### (Recurrent Neural Network Language Model)

- NNLM을 Recurrent Neural Network의 형태로 변형 한 것.(Short-term 메모리 역할 수행)
- 단어 정해줄 필요 없음, 단어를 순차적으로 입력해 주는 방식으로 학습 진행
- 굉장히 많은 데이터 학습이 필요.



# 워드 임베딩: Word2Vec

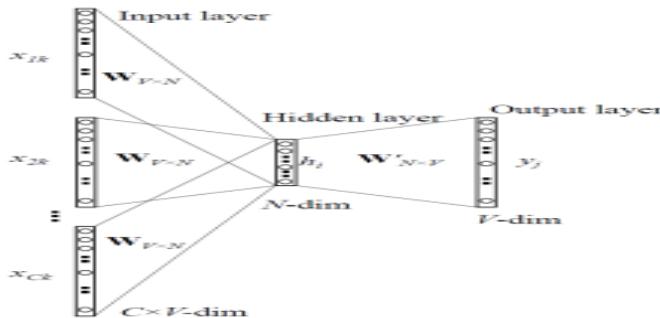


## Word2Vec

- 2013년에 구글에서 발표한 Word-Embedding 방식
- 기존의 계산량을 획기적으로 감소시켜서 몇배 빠른 학습이 가능.
- 가장 많이 사용하고 있는 Word-Embedding 모델

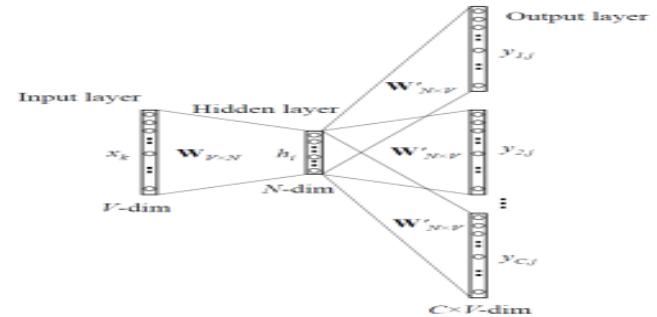
### CBOW(Continuous Bag-of-Words) 모델

- 주어진 단어에 대해 앞/뒤로의 단어를 input으로 사용하여 주어진 단어를 맞추는 네트워크. 모델은 Input Layer, Projection Layer, Output Layer로 이루어짐.



### Skip-gram 모델

- 주어진 단어 하나를 가지고 주변 여러 단어들에 대해서 확률적으로 유추하는 것
- CBOW가 더 빠른 연산량을 지니지만, 다소 더 나은 결과 값을 보여주고 있음.



# 워드 임베딩: GloVe



## GloVe(Global Vectors for Word Representation)

- 2014년 미국 스탠포드 대학 연구팀이 개발한 단어 임베딩 방법론
- LSA(Latent Semantic Analysis)와 Word2Vec의 단점을 보완하고자 개발한 기술
- 동시등장확률 개념 등장(probability of co-occurrence)
  - 학습말뭉치에서 동시에 같이 등장한 단어의 빈도를 각각 세어서 전체 말뭉치의 단어 개수로 나눈 결과 값.
- 특정 문맥 단어가 주어졌을 때, 임베딩된 두 단어 벡터의 내적이 두 단어의 동시등장확률간 비율이 되게끔 임베딩.

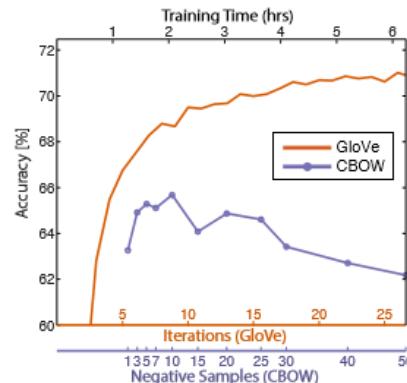
### LSA(Latent Semantic Analysis)

- Global matrix factorization 기술을 활용하여 말뭉치 전체의 통계적 정보를 활용.
- 단어 및 문서간 유사도/관계를 측정하기 어려운 단점 지님

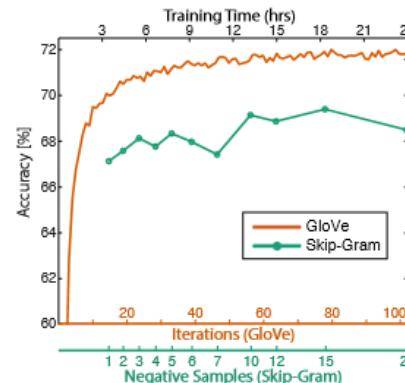


### Word2Vec

- 임베딩된 단어속에서 유사도 측정에 높은 성능 보임.
- 전체 정보 이용 대신 일부 단어 정보만을 활용하여서 추측.



(a) GloVe vs CBOW



(b) GloVe vs Skip-Gram

<출처 : GloVe, Stanford>



## 워드임베딩 기술(Word-Embedding)

### 유사도에 따른 분류 예시

- 각 단어의 벡터값에 따른 연관성이 높은 단어들을 나열.

FRANCE	JESUS	XBOX	REDDISH	SCRATCHED	MEGABITS
AUSTRIA	GOD	AMIGA	GREENISH	NAILED	OCTETS
BELGIUM	SATI	PLAYSTATION	BLUISH	SMASHED	MB/S
GERMANY	CHRIST	MSX	PINKISH	PUNCHED	BIT/S
ITALY	SATAN	IPOD	PURPLISH	POPPED	BAUD
GREECE	KALI	SEGA	BROWNISH	CRIMPED	CARATS
SWEDEN	INDRA	PSNUMBER	GREYISH	SCRAPED	KBIT/S
NORWAY	VISHNU	HD	GRAYISH	SCREWED	MEGAHERTZ
EUROPE	ANANDA	DREAMCAST	WHITISH	SECTIONED	MEGAPIXELS
HUNGARY	PARVATI	GEFORCE	SILVERY	SLASHED	GBIT/S
SWITZERLAND	GRACE	CAPCOM	YELLOWISH	RIPPED	AMPERES

# 워드 임베딩: 유사도에 따른 분류 예시

## 워드임베딩 기술(Word-Embedding)

## 유사도에 따른 분류 예시

- 각 단어의 벡터값에 따른 연관성이 높은 단어들을 나열.

#### □ 서울시/nnp 에 따른 워드임베딩 결과

유사 키워드	품사	유사도
대전시	nnp	0.872288745481
대구시	nnp	0.863960947392
광주시	nnp	0.861110910976
서대문구청	nnp	0.854604645319
부산시	nnp	0.849407190551
서울특별시	nnp	0.84725031642
관악구	nnp	0.845485691971
인천시교육청	nnp	0.845327081104
울산시	nnp	0.84488487431
관악구청	nnp	0.839168328143
서귀포시청	nnp	0.834813048457
충남도	nnp	0.831908366916
광명시	nnp	0.828996607731
서초구청	nnp	0.82861986766
노원구	nnp	0.827777197305
아산시	nnp	0.827628197551
천안시	nnp	0.827323302249
충주시	nnp	0.824166605091
금천구	nnp	0.823213377638



# [인공지능 기술의 이해]

- ① 4차 산업혁명과 인공지능
- ② 인공지능 개론
- ③ 알고리즘 설명 및 활용



## 기계학습(Machine Learning)이란?

기계학습 모델은 사용자가 직접 학습. 기계가 일일이 코드로 명시하지 않은 동작을 **데이터로부터 학습**하여 실행할 수 있도록 하는 알고리즘을 개발하는 연구 분야.

### 기계학습의 의미

- 학습 : 어떠한 지식이나 기술 등을 배우고 익히는 일
- 기계학습 : 의미 그대로, 기계가 학습하는 일.

### 기계 학습의 속성

- 데이터를 분석을 통한 새로운 지식 발견
- 컴퓨터가 자체적으로 발견 할 수 있도록 하는 알고리즘 학습
- 지식 자체가 아니라 지식을 찾는 알고리즘
- 알고리즘을 통한 간단한 예측 가능

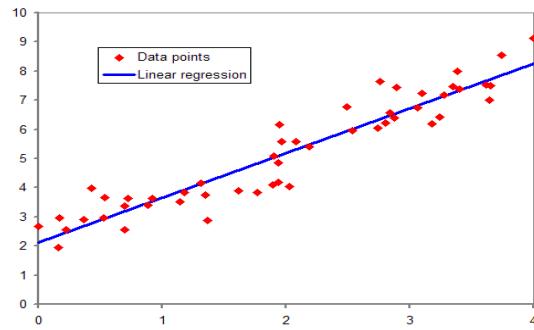
# 기계학습(Machine Learning)의 3가지 학습?

## (1) 지도 학습(Supervised learning)

- 사람이 교사로써 각각의 입력(x)에 대해 레이블(y)를 달아놓은 데이터를 컴퓨터가 학습할 수 있도록 하는 방법.
- 레이블 y를 잘 예측하는 것이 목적
- 대표적인 방법 : 회귀분석, 분류분석, 의사결정나무, 서포트 벡터머신 등

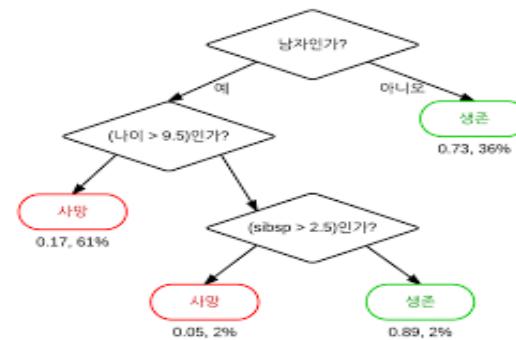
### 회귀분석모델 (Regression Model )

- 종속변수와 독립변수 사이의 관계를 분석하기 위해서 사용하는 모델.
- 독립변수 개수에 따라서 단순회귀분석과 다중 회귀분석으로 나뉨.
- 데이터들 사이에서 회귀식으로 특징 추출.



### 의사결정나무모델(Decision Tree Model)

- 각 변수에 따른 예측 가능한 규칙 및 패턴들을 나뭇가지 모양으로 형성한 것.
- 데이터 분류 결정 나무 및 회귀 결정 나무 존재.
- 분류하는데에 많이 사용.



# 기계학습- (1) 지도 학습 (Supervised Learning)

## 나이브 베이즈 모델 (Naïve Bayes Model)

- 조건부 확률 식을 응용하여서 적용한 모델. (Bayes's rule의 적용을 기본으로 함.)
- 데이터 특징에 따른 확률들을 계산하여 분류. 모든 특성값은 서로 독립임을 가정.(특성들 사이에 연관성 없음)
- 텍스트 분류에 많이 사용.(스팸문자 분류 등)
- Laplace Smoothing, Log Conversion을 통해서 단점을 극복 하고 있음.

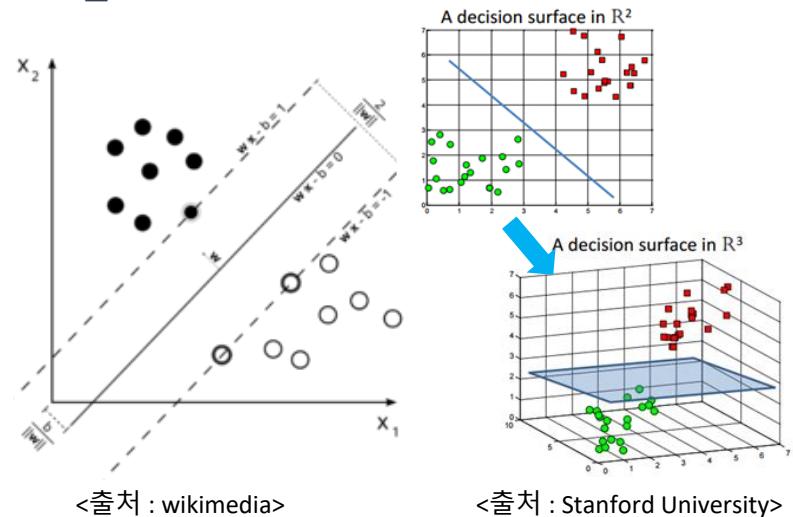
$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}$$

$$P(B|A) = \frac{P(A*B)}{P(A)}$$

$$\begin{aligned} P(X|C_i) &= \prod_{k=1}^n P(x_k|C_i) \\ &= P(x_1|C_i) \times P(x_2|C_i) \times \dots \times P(x_n|C_i) \end{aligned}$$

## 서포트 벡터머신 (Support Vector Machine)

- 분류 및 회귀를 위한 비확률적 이진 선형 분류 모델.
- 다수의 데이터를 분류 하기 위한 최적의 초평면 (Hyperplane)을 찾아내는 머신러닝 알고리즘.
- 선형 분류 및 비선형 분류 가능.
- 마진/서포트벡터/커널 개념 필요.
- 딥러닝 이전에 가장 널리 쓰인 기계학습 알고리즘.



# 기계학습- (1) 지도 학습 (Supervised Learning)

## 지도 학습(Supervised learning)기법에 따른 장단점.

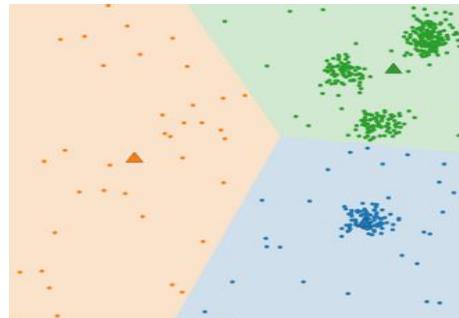
	Regression	Decision Tree	Naïve Bayes	Support Vector Machine
장점	<ul style="list-style-type: none"><li>1) 결과를 통해 유효한 정보를 획득할 수 있음.</li><li>2) 필요없는 변수 선택을 통해 모델의 안정성을 높일 수 있음.</li></ul>	<ul style="list-style-type: none"><li>1) 분석가가 결과를 쉽게 이해하고 설명 가능.</li><li>2) 확률에 따라서 정확히 분별 가능해짐.</li></ul>	<ul style="list-style-type: none"><li>1) 기준에 따라 합리적인 의사결정 형성 가능.</li><li>2) 계산이 간단함</li><li>3) 복잡한 상황에서 사용 가능</li></ul>	<ul style="list-style-type: none"><li>1) 사용자가 설정해야 하는 값이 적음.</li><li>2) 높은 정확도를 자랑함.</li></ul>
단점	<ul style="list-style-type: none"><li>1) 사전에 결측치 처리 필요.</li><li>2) 변수 간 교호작용의 유무 및 비선형 여부를 파악해야 함.</li></ul>	<ul style="list-style-type: none"><li>1) 확실한 답만 가능, 복잡할 수록 정확도 내려감. (트리 복잡성 증가)</li><li>2) 구분이 확실한 데이터만 가능함.</li></ul>	<ul style="list-style-type: none"><li>1) 좋은 성과를 얻기 위해 많은 수의 레코드가 필요</li><li>2) 편향된 결과를 획득할 가능성이 존재.</li></ul>	<ul style="list-style-type: none"><li>1) 최적의 커널을 자동으로 설정하기가 어려움.</li><li>2) 고차원에 따라서 연산량이 증가하게 되어 시간이 오래 걸림.</li></ul>

# 기계학습- (2) 비지도 학습 (Unsupervised Learning)

- 사람 없이 컴퓨터가 스스로 레이블이 없는 데이터에 대해 학습하는 것.
- $y$ 값 없이  $x$ 값만을 이용하여 학습하는 것을 의미.
- 실제로 기준으로 분류한다는 개념이 아닌 확률적으로 비슷한 것끼리 묶어주는 의미.
- 대표적인 방법 : Clustering, K-means 알고리즘, 밀도추정 등

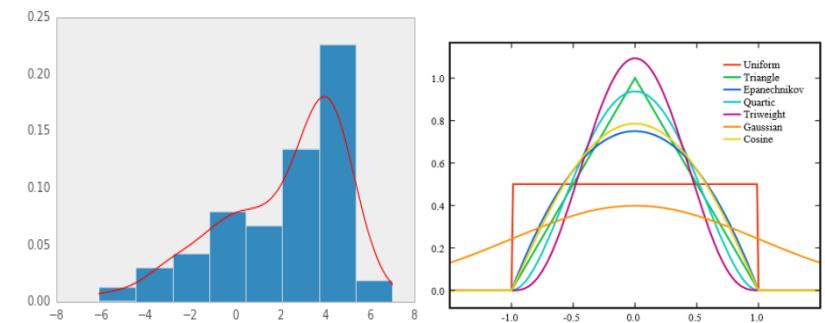
## K-means 알고리즘

- 데이터를 k개의 군집으로 묶어주는 알고리즘.
- 각 군집 간 거리 차이의 분산 값을 최소화 하는 방식으로 동작.
- 초기 평균값을 데이터 중에 랜덤으로 k개 뽑아 실시, 클러스터내의 무게중심을 최적화.
- 용도 : 주제별 분류, 네트워크 유해 트래픽 탐지, SNS 유사사용자 군집화, 유사고객집단추출 등



## 밀도추정 (Density Estimation)

- 데이터 분포로부터 특정 변수가 가지는 확률 분포의 특성 추정하는 것.
- 데이터와 변수의 관계를 파악하기 위한 방법.
- 대표적으로 히스토그램 밀도추정 방법과 커널 함수 밀도 추정 방법이 존재.



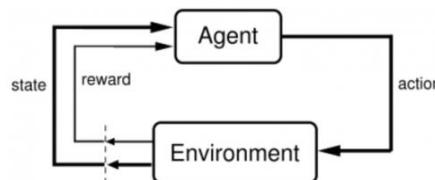
<출처 : wikipedia>

# 기계학습- (3) 강화 학습 (Reinforcement Learning)

- 현재의 상태(State)에서 어떤 행동(Action)을 취하는 것이 최적인지를 학습하는 방법.
- 행동을 취할 때마다 외부 환경에서 보상(Reward)가 주어지는데, 이러한 보상을 최대화 하는 방향으로 학습이 진행.
- 대표적인 방법 : Deep-Q learning, Genetic Algorithm 등

## Deep-Q Network(DQN)

- Deep learning과 Q-learning 방식을 결합한 것.
- 학습의 보상은 게임 점수, 컨트롤러 조작을 학습 하여 점수 향상.
- 게임 규칙 부여 하지 않은 상태로 게임 반복
- 입력과 출력을 통해서 자신이 알아서 학습해 나 가는 것.



Q-learning 알고리즘.  
$$Q(s,a) = (s,a) + \alpha(R + \gamma \max_{a'} Q(s',a') - Q(s,a))$$

<출처 : The Agent-Environment Interface>

## 유전자 알고리즘(Genetic Algorithm)

- '우수한 개체 = 좋은 해답'이라고 가정한 진화 기법을 이용해서 최적의 답을 찾고자 하는 알고리즘.
- 세대가 지날 수록 진화를 거듭.
- 염색체/유전자/교차/돌연변이/자손의 유전자용어 사용.

### 유전자 알고리즘 설명

- 1) 초기 세대 생성
  - 2) 초기 세대 적합도 평가
  - 3) 초기 세대 교배 및 돌연변이 (해당 과정으로 다음 세대가 생성된다)
  - 4) 다음 세대 적합도 평가
- 5-1) 목표 적합도를 만족하는 염색체가 존재하면 알고리즘을 종료
- 5-2) 목표 적합도를 만족하는 염색체가 없다면 다음 과정으로 진행
- 6) 현재 세대 교배 및 돌연변이 (해당 과정으로 다음 세대가 생성된다)

<출처 : white whale>



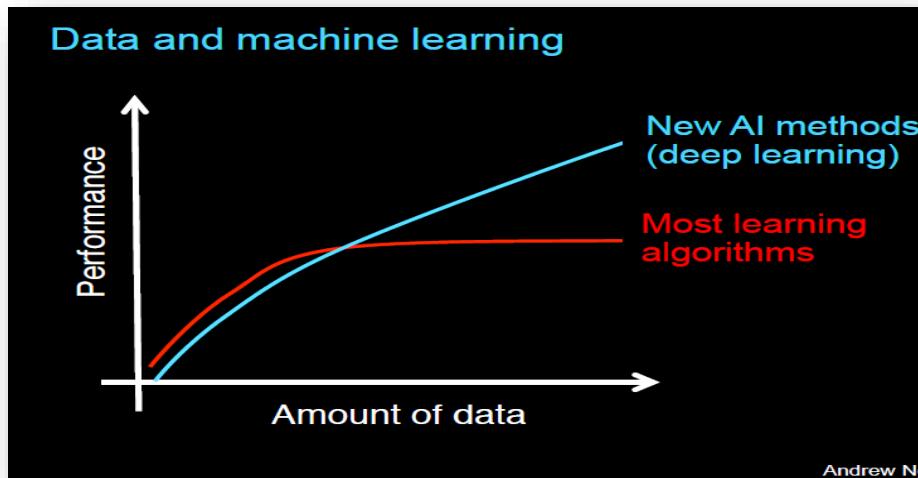
## 딥러닝의 부활

- 2006년 Science 저널에 발표된 토론토 대학의 Hinton 교수의 논문 이후로 활발하게 관심 받기 시작.
- MIT에서 선정한 2013년을 빛낼 10대 혁신 기술 중 하나로 선정
- 가트너가 2019년 세계 IT 시장 3대 주요 트렌드에 속한 주요 기술로 지정.

## 딥러닝 대표 학자

- Geoffrey Hinton : 캐나다 토론토 대학교 교수겸 구글 연구 소장
- Yann LeCun : 뉴욕 대학교 교수 겸 페이스북 인공지능 연구소장
- Andrew Ng : 스탠포드 대학교 교수 전 바이두 인공지능 연구소 소장.

## 딥러닝의 성능



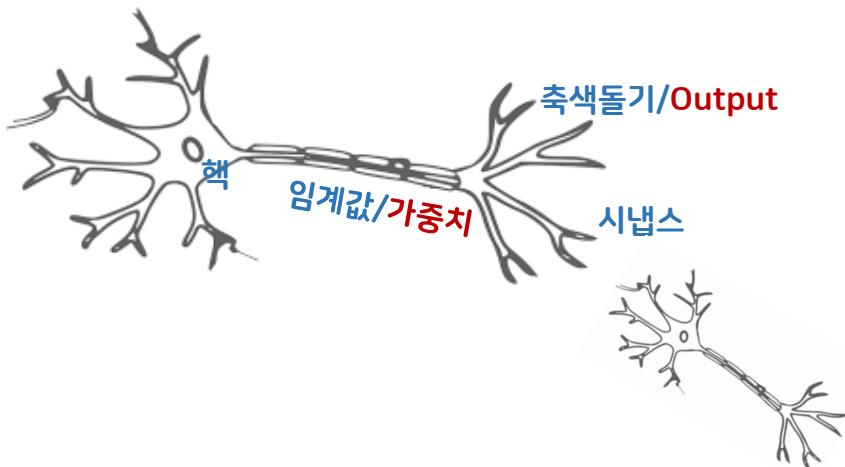
# 인공신경망의 개념



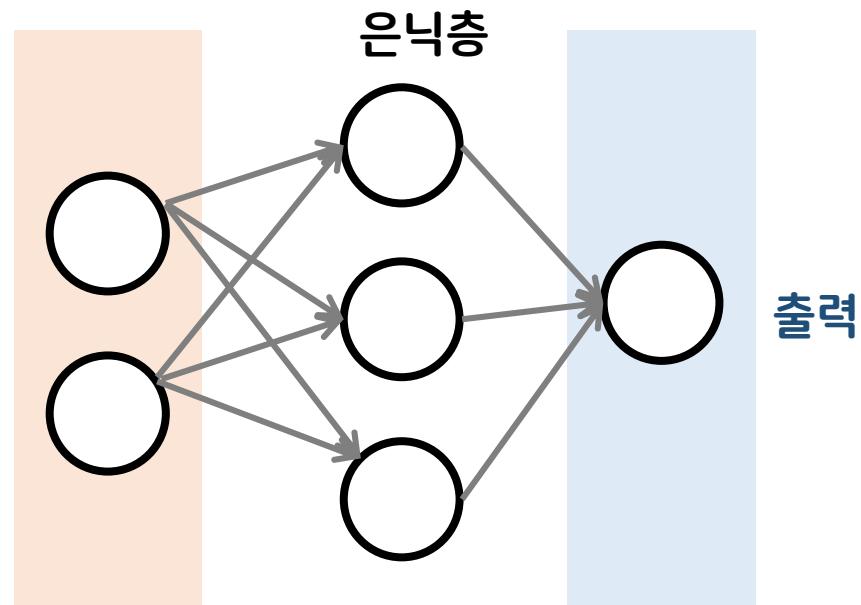
## 인공신경망 (Artificial Neural Network)

- 인간의 신경망 원리 및 구조로부터 영감을 얻어 형성한 기계학습 알고리즘.
- 뉴런은 수상돌기에서 신호를 받아들여 시냅스를 통해서 축색돌기로 전달하는 구조. 전달하기 위해서 시냅스에 임계값이상의 전기 신호 필요.
- 동일한 원리로, input/output/threshold의 값으로 표현
- 모델 학습 목표 : 각 노드의 적정 가중치 값을 알아내는 것.  
= 시냅스의 적절한 임계값을 찾는 것.

수상돌기/Input



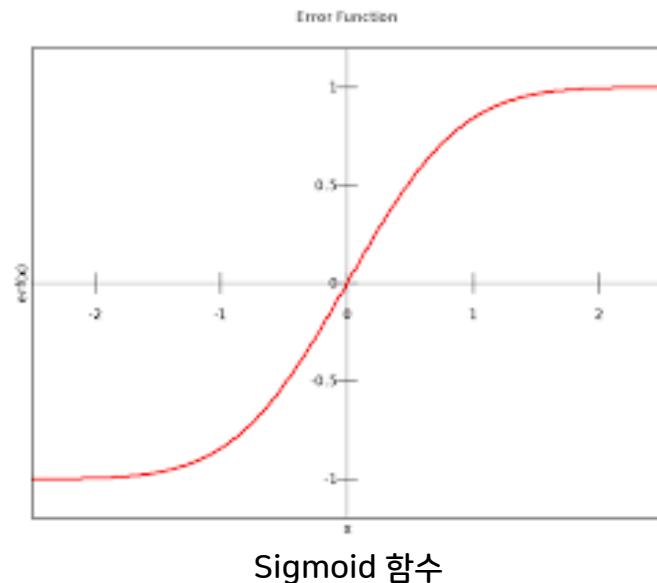
입력



# 인공신경망의 문제점



- 학습 과정 속에서 Parameter들의 최적의 값을 찾기 어려움
  - ✓ 기울기 값이 뒤로 전달 될수록 점점 작아짐.(vanishing gradient problem)
  - ✓ 부분적으로 에러율이 최저인 부분을 일반화 시키는 오류(Local minima 현상)
  - ✓ Sigmoid 사용으로 인하여 학습에 한계.
- 학습 데이터(Training data)에 따른 과적합(Overfitting) 문제
  - ✓ Data가 많지 않은 경우에 발생.
  - ✓ 학습데이터에만 특화되어 있는 모델이 형성.
- 느린 학습 시간
  - ✓ 은닉층이 증가될수록, 연산량이 급격히 증가.
  - ✓ 하드웨어의 한계.



# 인공신경망의 해결책



- 과적합(Overfitting) 및 Local Minima 해결
  - ✓ 사전훈련(pre-training)을 통해 과적합(Overfitting)을 방지할 수 있는 initialize point 알고리즘 개선
  - ✓ Rectified Linear Unit(ReLU), Drop-out 등
- 느린 학습 시간 해결
  - ✓ 하드웨어의 발전 : 그래픽 카드 발전 (GPGPU, CUDA 등)
  - ✓ 최근 인공지능 전용칩 개발
- 빅데이터의 출현
  - ✓ 대량의 쏟아져 나오는 데이터를 이용

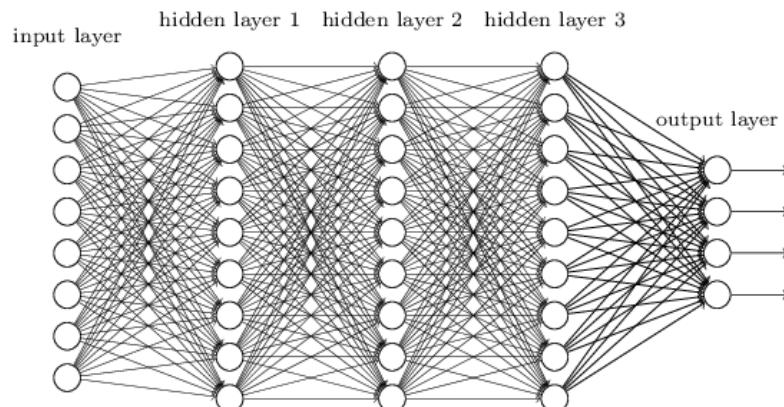
# 단층 신경망 vs 다층 신경망



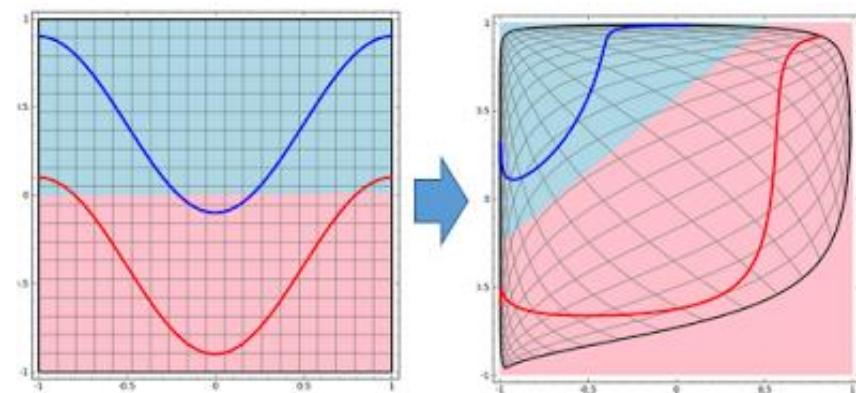
## Deep learning과 DNN (Deep Neural Network)

- 딥러닝은 2개이상의 은닉층을 지닌 학습 방법.
- DNN은 기존 인공신경망 구조에서 Deep Learning기법을 도입하여 한층 고차원의 연산이 가능한 기계학습 방법.
- 공간을 왜곡하고 데이터를 구분 짓는 과정을 반복하여 최적의 구분선을 도출.
- 목적을 위해 많은 데이터와 많은 반복 과정이 필요함.
- 최근에 들어서 Deep Learning과 DNN을 혼용하여 사용함.
- 역전파(Back-propagation)과 사전학습(Pre-training)기법이 도입되면서 전성기.

Deep neural network



<사진출처 : Wikipedia>



<사진출처 : [colah's blog](#)>

# 단층 신경망 vs 다층 신경망



## DNN의 응용

- 딥 러닝은 인간의 두뇌가 수많은 데이터 속에서 패턴을 발견 후, 사물을 구분하는 정보처리 방식을 모방해 컴퓨터가 사물을 분별하도록 기계를 학습시킴.
- 딥 러닝 기술을 적용하면 사람이 모든 판단 기준을 정해주지 않아도 컴퓨터가 스스로 인지·추론·판단할 수 있게 된다.
- 음성·이미지 인식과 사진 분석, 텍스트 인식 등에 광범위하게 활용됨.
- 대표적인 DNN 응용 알고리즘 : CNN,RNN,LSTM,GRU 등

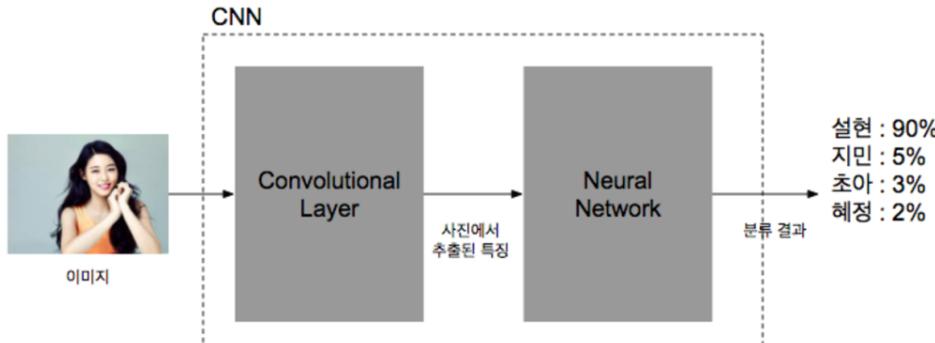


- ✓ 손으로 작성되어진 우편번호를 인식하는 Deep Neural Network에서 시작(Lechan at el., 1989)
- ✓ imageNet이라는 대회에서 월등한 성능으로 1등을 차지함(Alex Krizhevsky at el., 2012)
- ✓ 대량 어휘 음성인식을 통한 딥러닝으로의 성공적인 확장(Li Deng at el., 2009)
- ✓ 다양한 자연어처리를 위해 딥러닝 적용(R. Collobert, 2011)

# 합성곱신경망(CNN : Convolutional Neural Network)

## CNN 정의

- 인간의 시신경 구조에서 모방한 신경망 구조.
- 기존의 구조는 데이터에서 지식을 추출하는 형태로 학습이 이루어졌으나, CNN은 데이터에서 피쳐를 추출하여 이를 통해 패턴 인식을 하는 구조.
- 입력 데이터로부터 고수준의 추상화된 정보를 추출하거나 또는 추출된 정보를 통해 새로운 정보 생성.(질 높은 특징 추출 기술이 필요)
- 파라미터 개수를 효율적으로 줄여주어, 전체 모델의 복잡성 감소.



<출처 : 조대협 (<http://bcho.tistory.com>)>



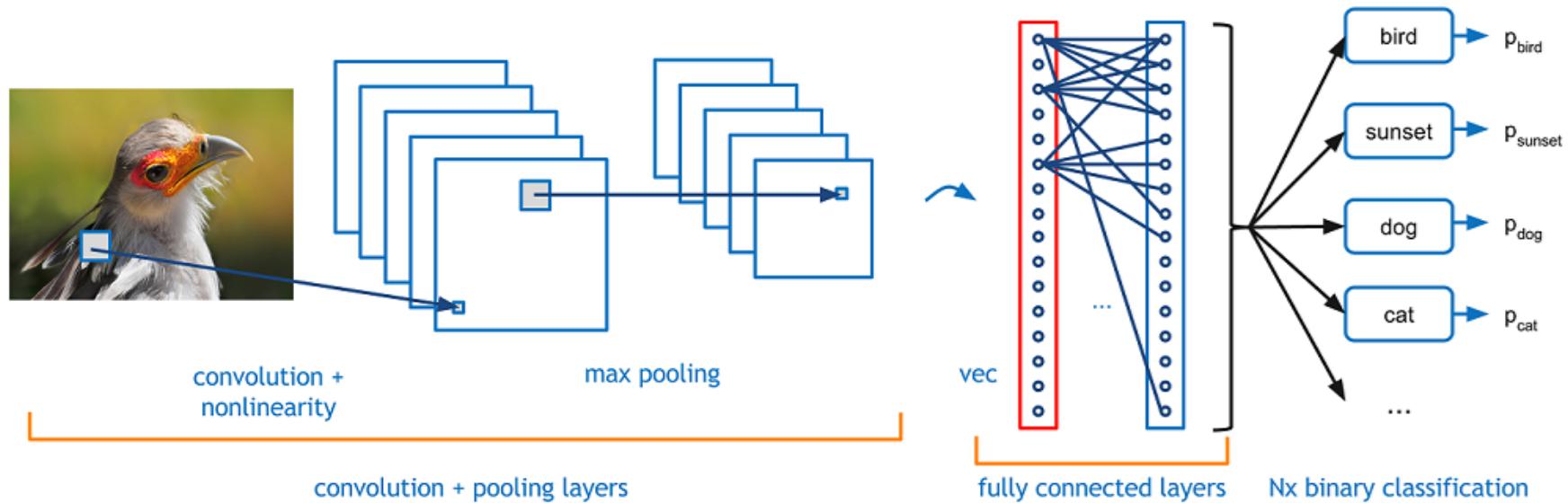
<출처: M. Zeiler>

# 합성곱신경망- CNN의 구조



## CNN의 구조

- 특징 추출을 위해서 convolution과정과 pooling과정을 실시함.
- 컨볼루션층 : 근접해 있는 노드들이 다음 층 노드와 연결되는 구조  
임의의 이미지에서 국소적인 부분을 추상화하는 역할
- 풀링층 : 이전층 노드의 국소적인부분을 통합하는 역할  
(맥스풀링 - 최댓값을 취하는 처리)



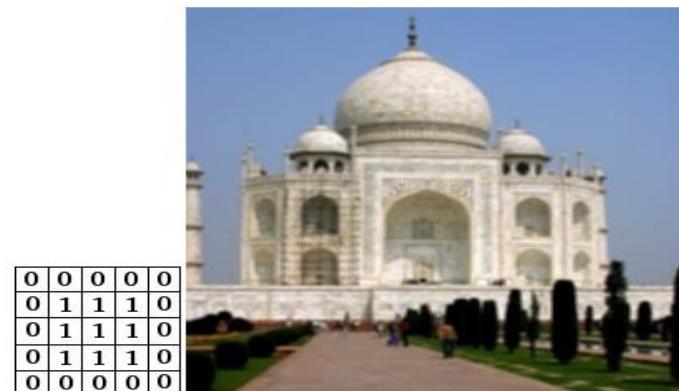
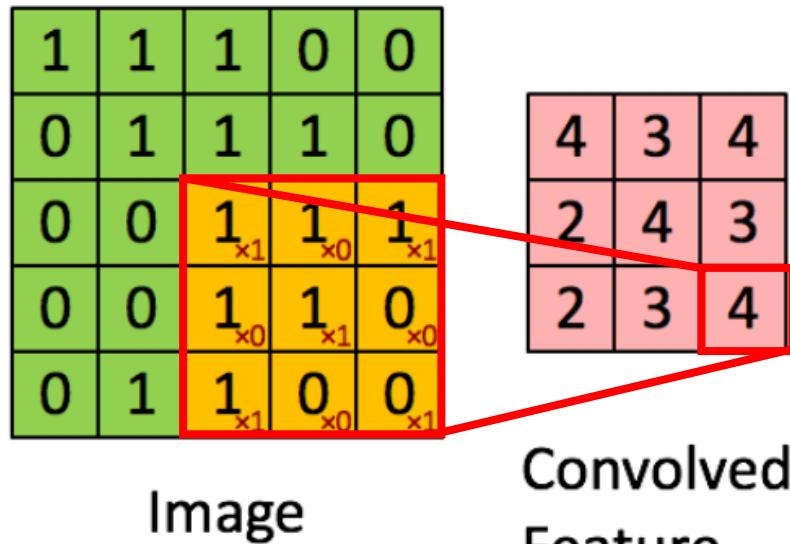
<출처: <https://adeshpande3.github.io/adeshpande3.github.io/>>

# 합성곱신경망- Convolution 과정

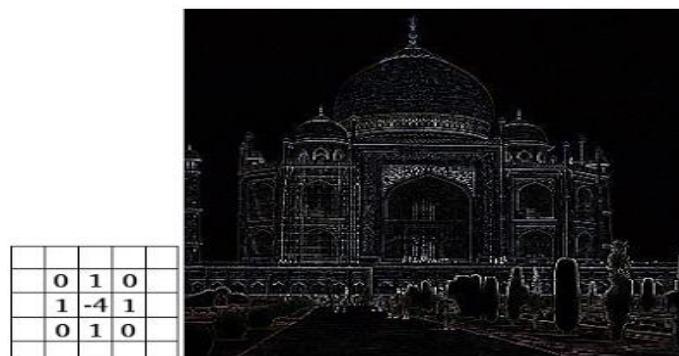


## Convolution 과정

- 특징을 추출하는 역할.
- 필터(Filter)와 액티베이션 함수(Activation 함수)로 이루어짐.
- 필터(Filter) : 특징이 데이터에 있는지 없는지 검출해주는 함수.
- 액티베이션 함수(Activation function) : 피쳐의 유무를 수치화 해주기 위해서 비선형 값으로 바꿔주는 과정.



0	0	0	0	0
0	1	1	1	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0



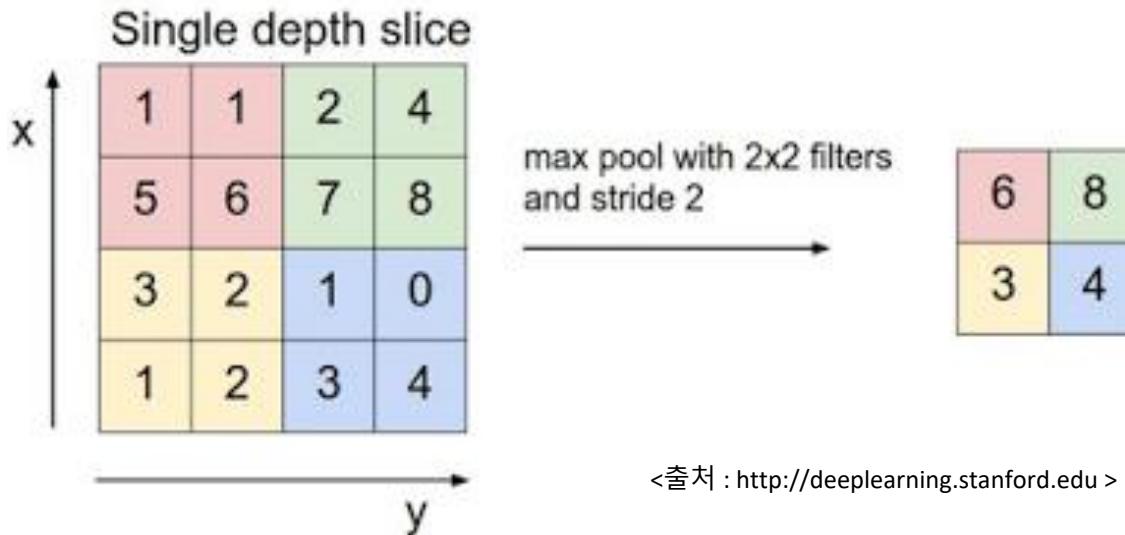
<출처 : <http://deeplearning.stanford.edu> >

# 합성곱신경망- Pooling 과정



## Pooling 과정

- 단순히 사이즈를 줄이는 과정.
- 여러 장점을 제공.
  - ✓ 미세한 부분 변화에 대해서 일관된 피쳐를 제공(찌그러진 차-보통 차)
  - ✓ 노이즈의 역할을 상쇄
- Sub-sampling(서브샘플링)이라고도 함.
  - ✓ Max-pooling, Average-pooling, L2-norm pooling 등

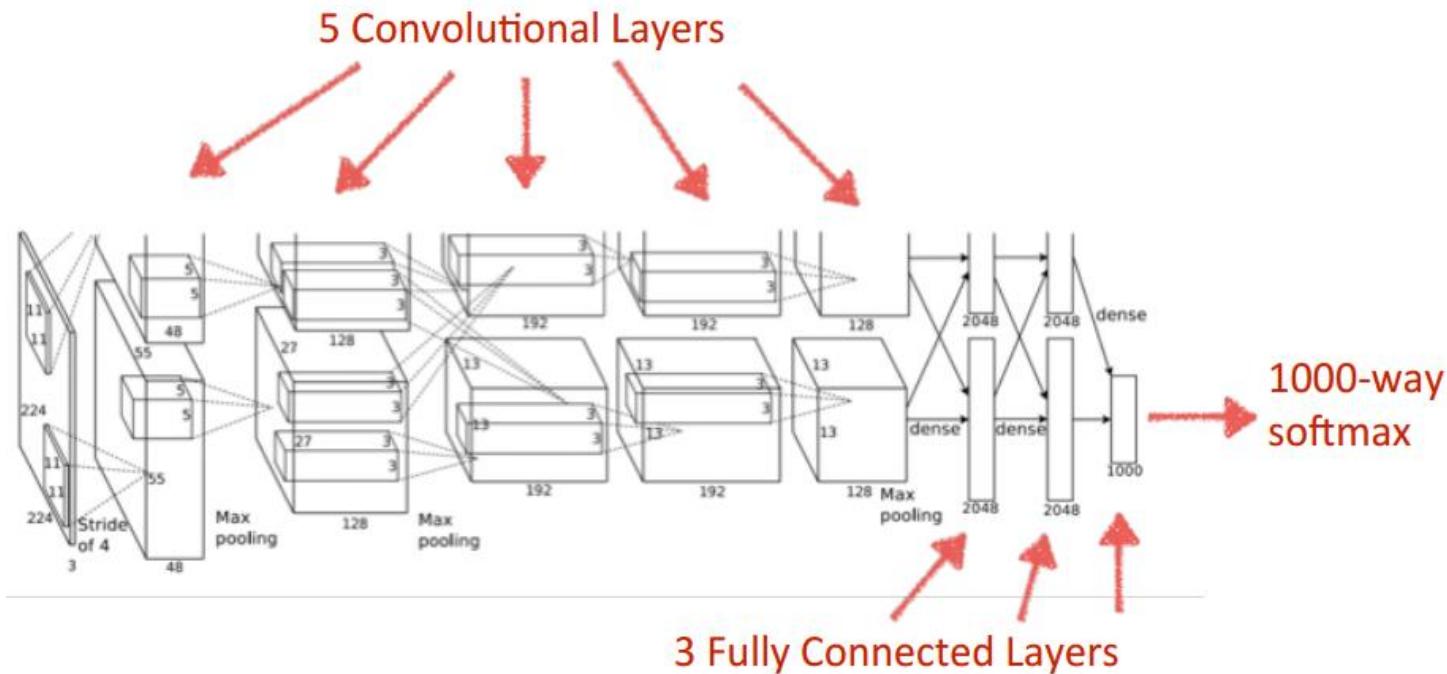


# 합성곱신경망- CNN의 다양한 구조



## CNN의 다양한 구조

- Convolution Layer와 Pooling Layer를 다양하게 구성하여서 실행.



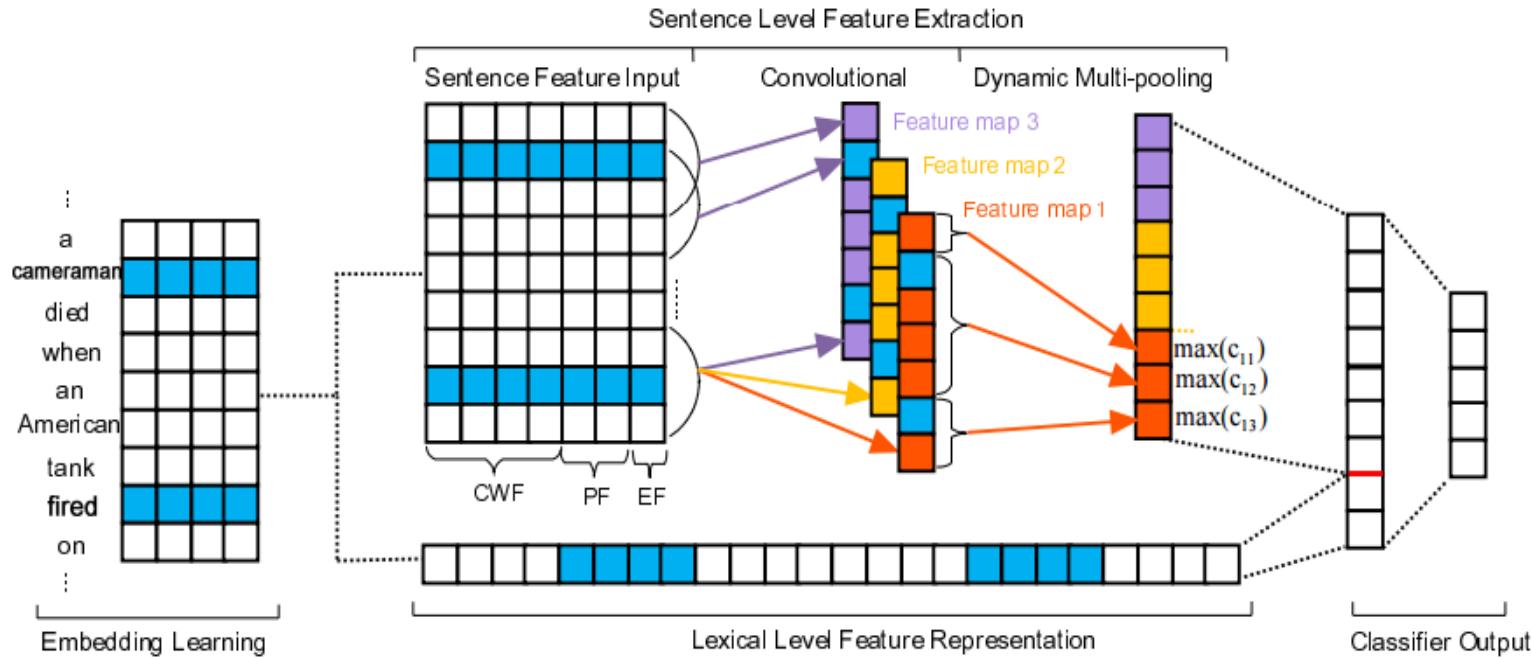
<출처: ImageNet Classification with Deep Convolutional Neural Networks(Alex Krizhevsky at el., 2012)  
modified: <http://vision.stanford.edu> by Tugce Tasici, Kyunghee Kim >

# 합성곱신경망- CNN의 활용



## CNN의 활용

- Information Extraction(정보 추출)
  - ✓ In Baghdad, a cameraman died when an American tank fired on the Palestine Hotel.



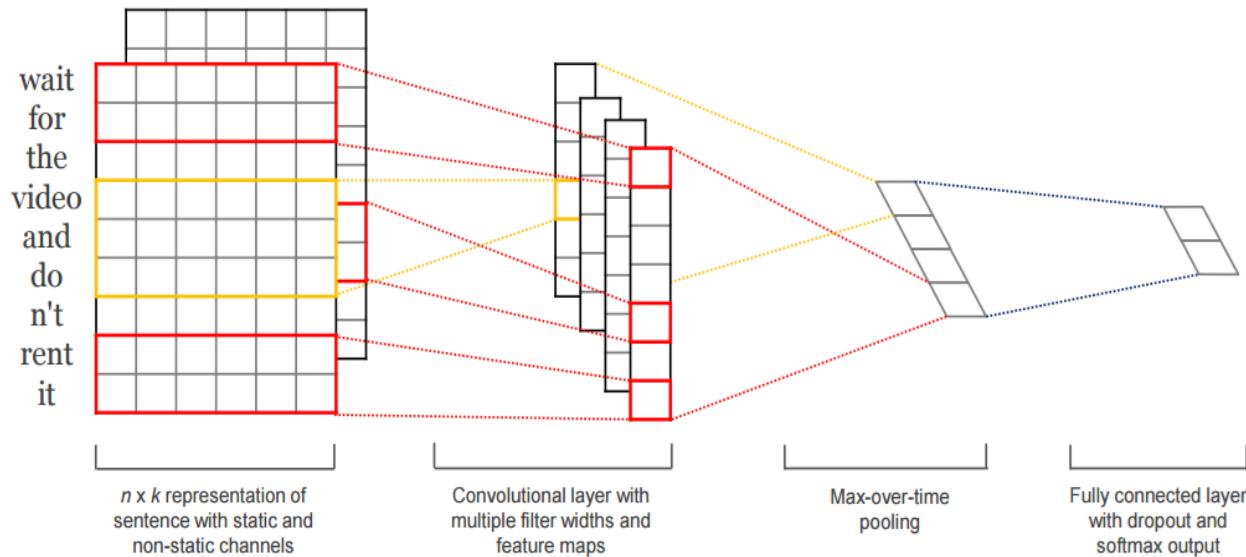
<출처: Event Extraction via Dynamic Multi-pooling Convolutional Neural Networks(Yubo Chen at el., 2015)>

# 합성곱신경망- CNN의 활용



## CNN의 활용

- Sentence Classification(문장 분류)  
✓ I regret to wait for the video and don't rent it

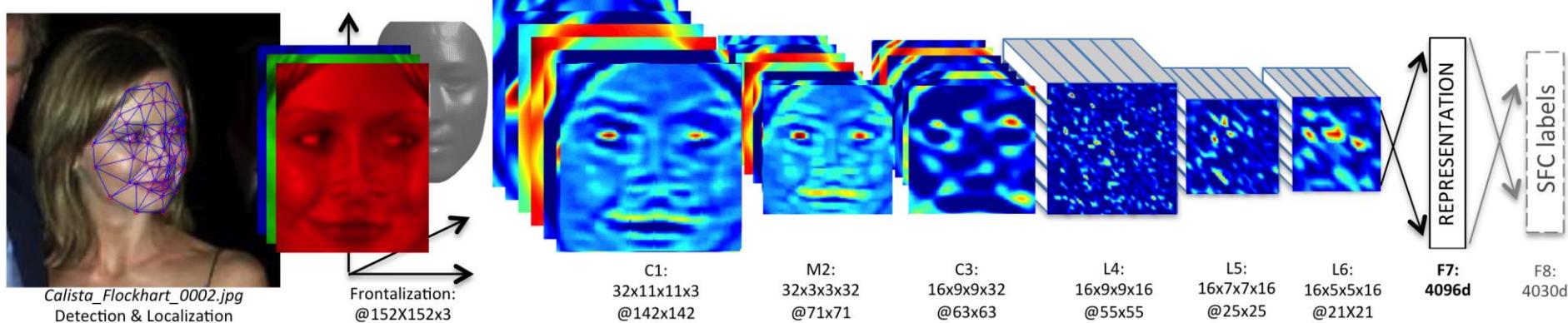


# 합성곱신경망- CNN의 활용



## CNN의 활용

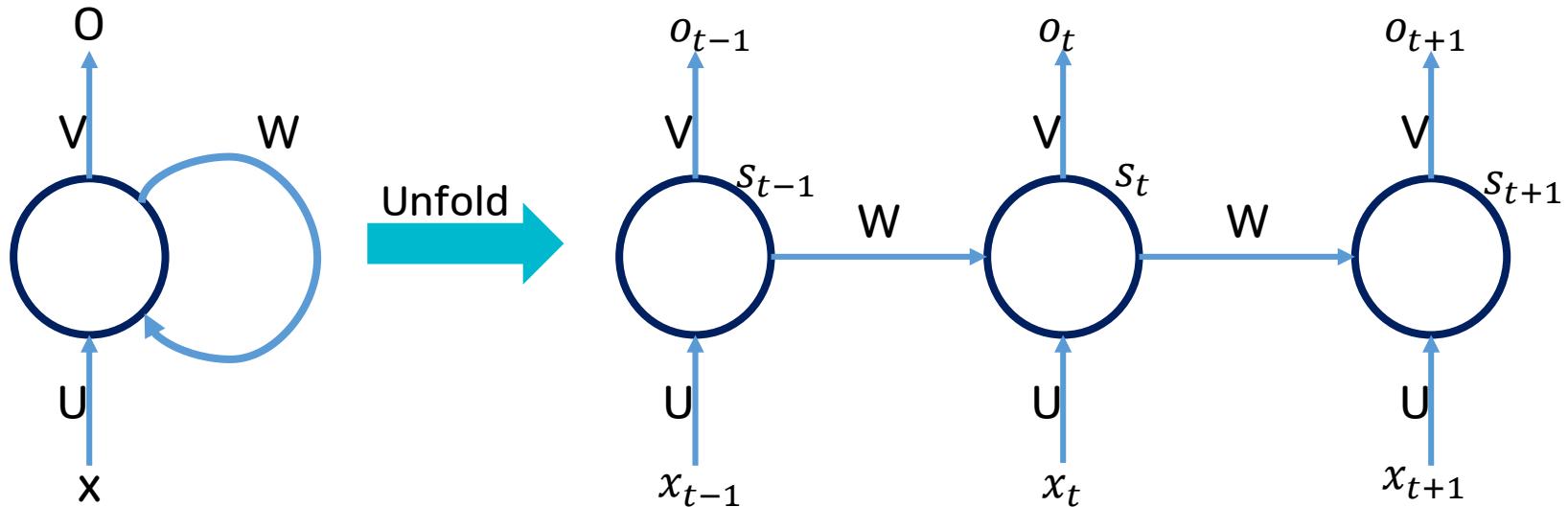
- Face Recognition(얼굴 인식)
  - ✓ Facebook에서 CNN(9 layers)을 기반으로 만든 얼굴 인식 알고리즘.
  - ✓ Human evaluation : 97.5%
  - ✓ DeepFace : 97.25%
- 실험 데이터셋 : 유튜브 얼굴 데이터셋



# 순환신경망(RNN : Recurrent Neural Network)

## RNN 정의

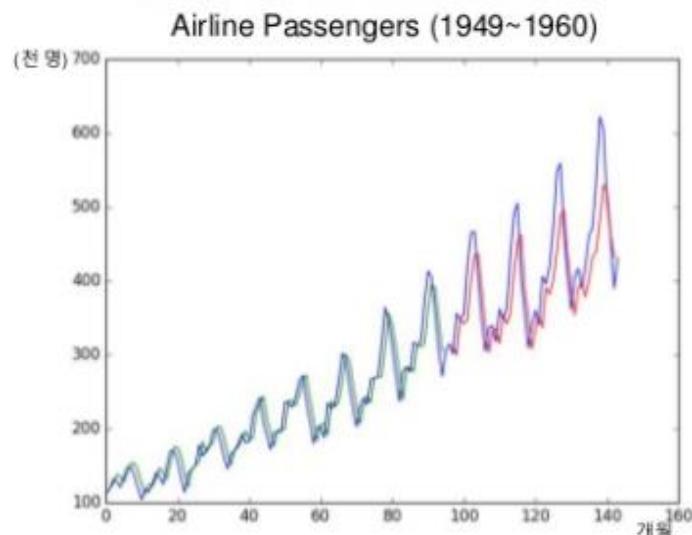
- 반복적/순차적 데이터(Sequential data) 학습에 특화 되어 있는 인공신경망의 한 종류
- 내부에 순환 구조로 이루어짐.
- 기존 알고리즘 문제 해결 : 지속성, 반복적 순차적 데이터 학습에 한계
  - ✓ 순환 구조로 인하여, 과거의 자신의 정보(가중치)를 기억하고 학습에 반영.
  - ✓ 이전 작업과 현재 작업과의 연결이 가능해짐.
- 텍스트 /음성 데이터 인식에 많이 쓰임.



# 순환신경망(RNN : Recurrent Neural Network)

## RNN 정의

- 반복적/순차적 데이터(Sequential data)  
시간이 진행됨에 따라서 변화가 있는 데이터.



Blue=Whole Dataset, Green=Training, Red=Predictions

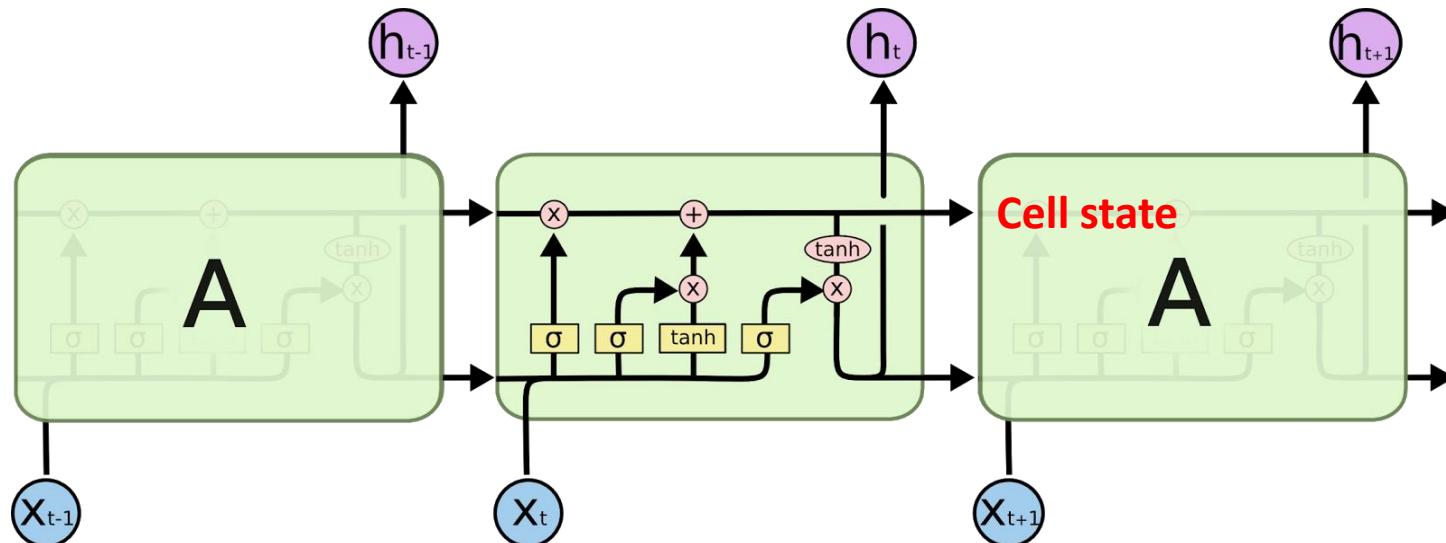
Musical Note  
"Plucked Guitar String"



# 순환신경망(RNN : Recurrent Neural Network)

## LSTM(Long Short Term Memory networks)

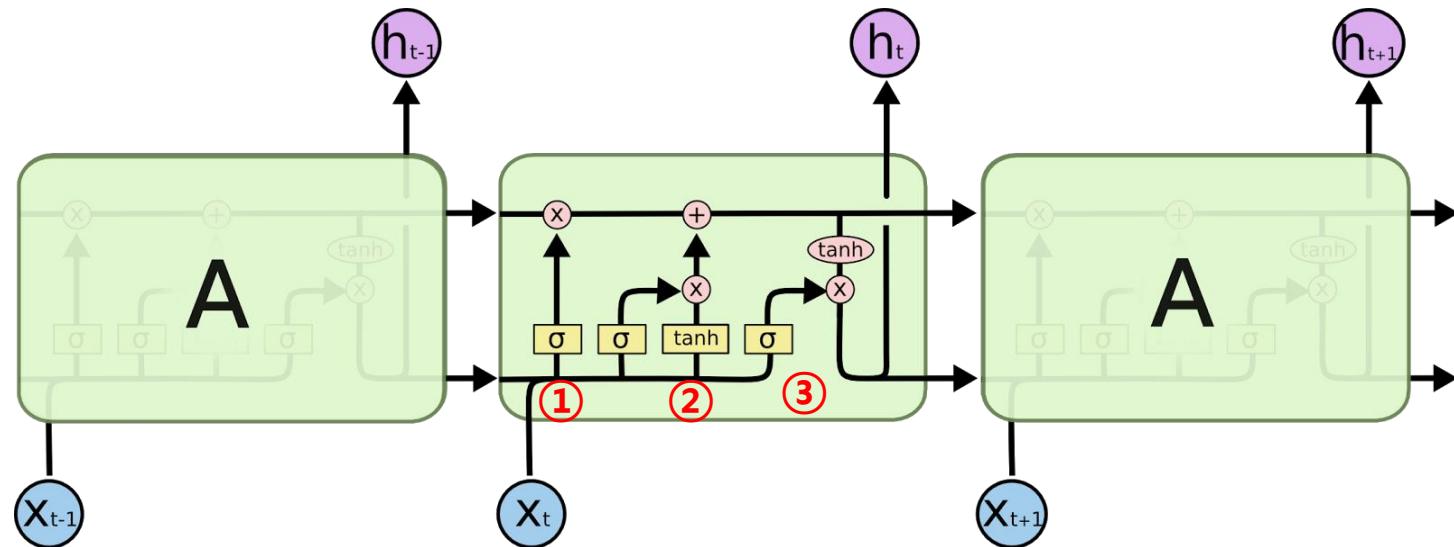
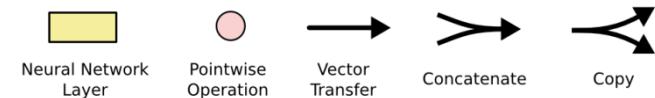
- RNN 알고리즘의 문제점을 해결 (The Problem of Long - Term Dependencies)
  - ✓ 기존 RNN은 근접한 과거 데이터에 대해서만 정보 유지.
  - ✓ 정보 사이에 거리가 멀 경우, RNN 학습능력 저하(Vanishing gradient problem)
- 히든 layer에 특별 기능을 하는 네 개의 층을 추가함.
- 핵심은 컨베이어 벨트처럼 정보가 흘러가는 셀 상태(Cell state).
- 셀 상태: 정보 추가 및 삭제 가능.



# 순환신경망(RNN : Recurrent Neural Network)

## LSTM(Long Short Term Memory networks)

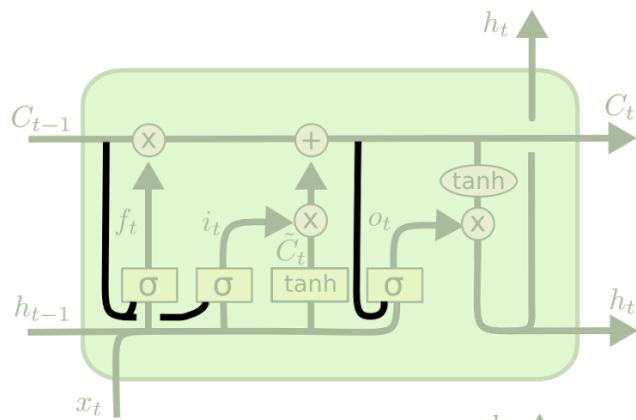
- 셀 상태 : 게이트 존재(활성화 함수와 요소별 곱셈 연산 구성)  
① 제거 단계 : 제거 게이트층(Forget Gate)에서 셀 상태에서 정보의 삭제 여부 결정.  
② 저장 단계 : 새로운 정보의 저장 여부 결정.
  - ✓ 입력게이트(Input Gate) : 갱신할 여부를 결정
  - ✓ tanh 층 : 갱신할 데이터에 대한 후보군 선정  
③ 출력 단계 : 필터링 된 버전의 결과를 출력여부 결정.



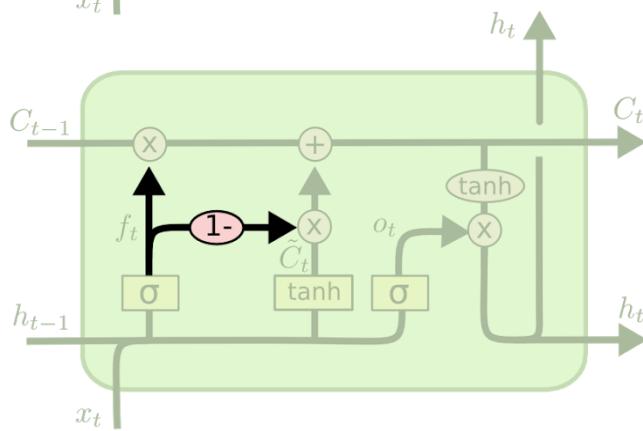
# 순환신경망(RNN : Recurrent Neural Network)

## LSTM(Long Short Term Memory networks) 응용

- 엿보기 구멍 (Peephole connections)
- Forget gate와 input gate 연결 : 입력하려고 할 때만 제거(불필요 연산 제거)



$$f_t = \sigma(W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f)$$
$$i_t = \sigma(W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i)$$
$$o_t = \sigma(W_o \cdot [C_t, h_{t-1}, x_t] + b_o)$$

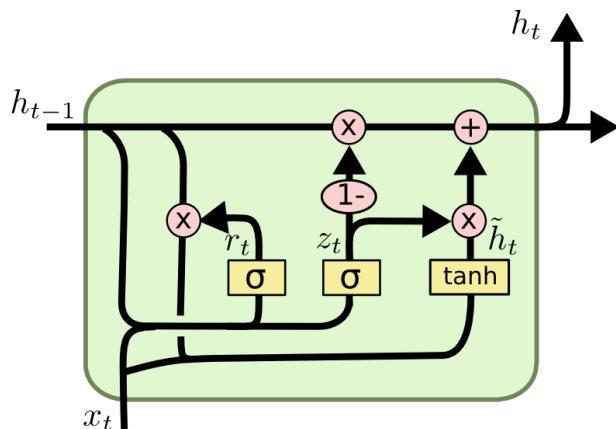


$$C_t = f_t * C_{t-1} + (1 - f_t) * \tilde{C}_t$$

# 순환신경망(RNN : Recurrent Neural Network)

## GRUs(Gated Recurrent Units)

- LSTM을 변형시킨 구조.
  - ✓ Gradient Vanishing 문제 극복 함. (LSTM과 유사)
  - ✓ 일부 게이트 생략 및 재구축
- Update gate와 reset gate로 나뉨.
  - ✓ Update gate: 과거 정보 반영 정도를 결정
  - ✓ Reset gate : 현시점 정보와 과거 시점 정보 반영 여부 결정



$$z_t = \sigma (W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma (W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh (W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

<출처: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>

# 순환신경망(RNN : Recurrent Neural Network)

## LSTM vs GRUs

	LSTM	GRUs
공통점	Vanishing gradient 문제 해결 Gate 개념 활용	
장점	각각 메모리 컨트롤이 가능 결과 값 컨트롤 가능	연산속도 빠름 메모리 덮어 씌워질 문제 없음
단점	메모리 덮어 씌워질 가능성 존재 연산속도 느림	메모리 컨트롤 불가 결과값 컨트롤 불가

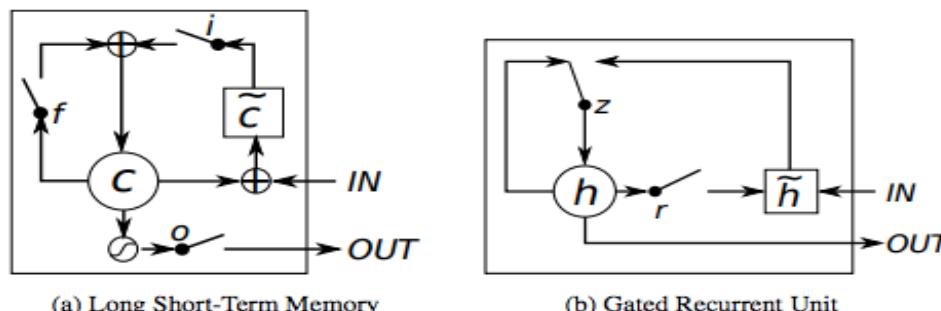
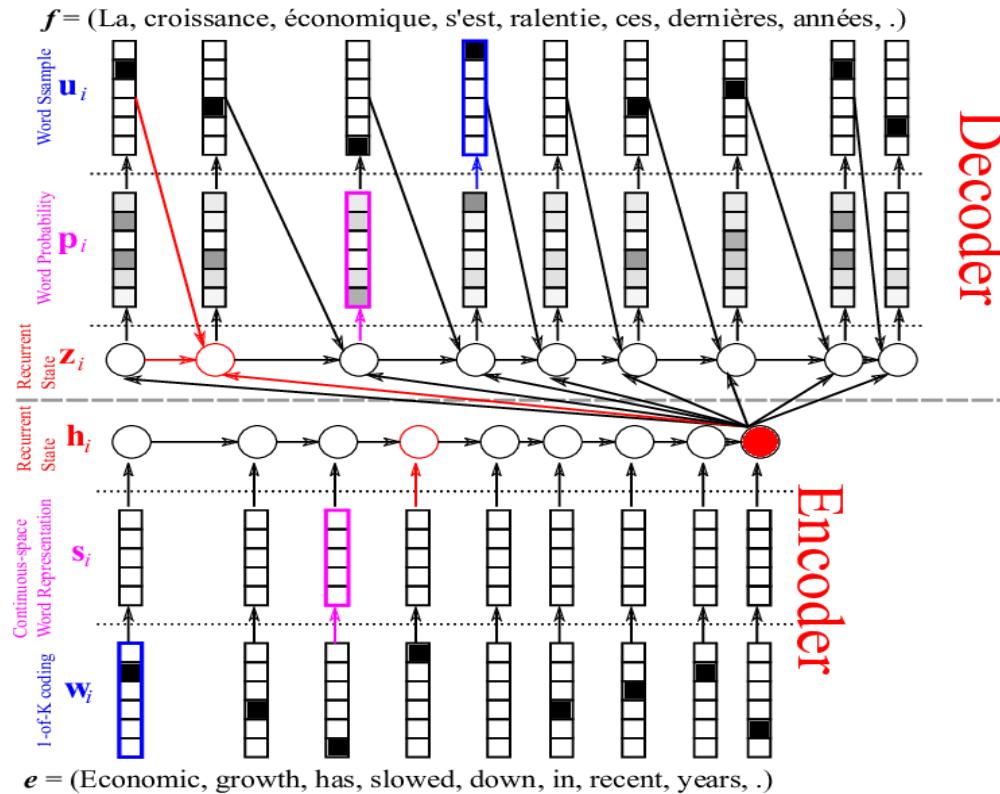


Figure 1: Illustration of (a) LSTM and (b) gated recurrent units. (a)  $i$ ,  $f$  and  $o$  are the input, forget and output gates, respectively.  $c$  and  $\tilde{c}$  denote the memory cell and the new memory cell content. (b)  $r$  and  $z$  are the reset and update gates, and  $h$  and  $\tilde{h}$  are the activation and the candidate activation.

# 순환신경망(RNN)의 활용



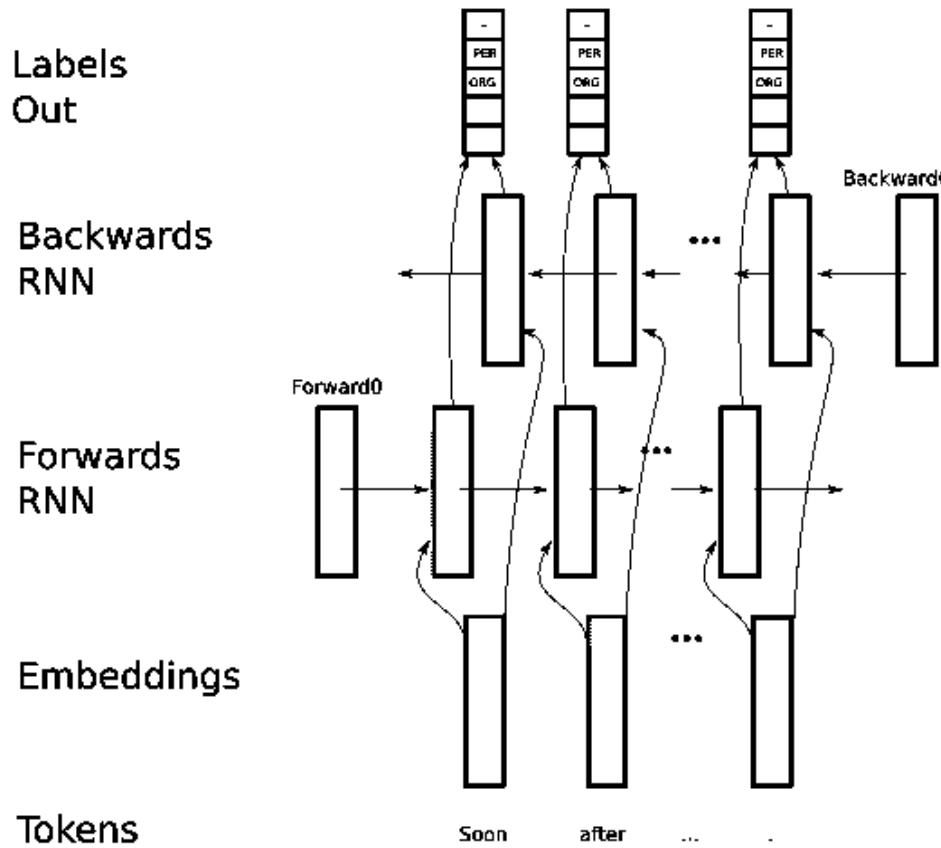
- Machine Translation(기계 번역)



# 순환신경망(RNN)의 활용



- Named Entity Recognition(개체명 인식)



# DNN의 활용 사례- 스팸문자 분류

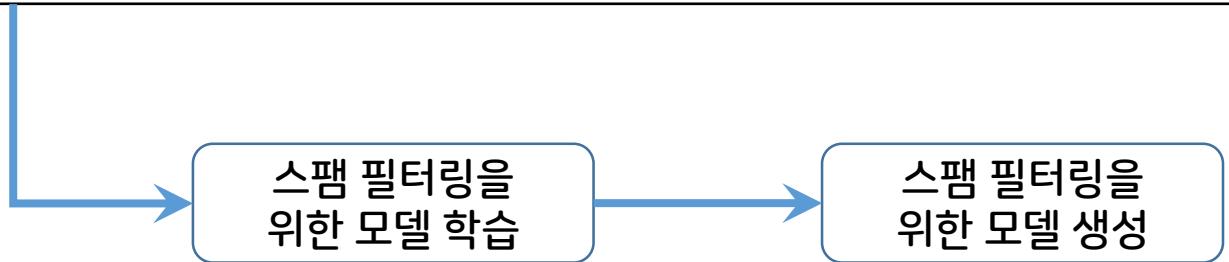


## 스팸문자 분류

기계학습 모델은 사용자가 직접 학습. 기계가 일일이 코드로 명시하지 않은 동작을 **데이터로부터 학습**하여 실행할 수 있도록 하는 알고리즘을 개발하는 연구 분야.

0: 일반문자, 1: 스팸문자

명절 잘보내세요 명절가입시 3만원받고 50만원 따기쉬움 도박.com	1
방금 너가 문자 보냈어?	0
안녕하세요 예전에 상담했던 김은영 입니다.	1
영국 아일랜드 어학연수자료 메일로 보내드렸습니다.	0
♥ br둑이,포카, brkrer.♥바다이야기.고래연타300만 24시대박	1
스카이 모임 공지 6월 26~27일 장소 계곡 참가여부 조사해요	0
저 민정인데요 예전에 통화한 오빠 맞죠? 사진보고 맞으면 연락주세요	1
소개팅시켜줘 외로워	0
오빠! 오늘밤이 외로워요 한잔하고싶어요 사진보고 연락	1



# DNN의 활용 사례- paypal



Payment 시장의 Innovative Leader PayPal은 사기방지를 위해 최첨단 분석 기법을 보유하고 있었으나, 고의적 사기를 막을 수 없었음.



이를 위해 Deep Learning 적용을 검토함

# DNN의 활용 사례- paypal



Complex Multilayered Network에 의한 데이터의 High-Level 패턴을 분석하는 Deep Learning 방식을 도입하여 사기방지 예측모델을 획기적으로 개선.

근본원인

기존의 탐지 방법과 달리 더 복잡한 체계로  
이루어진 모델을 제시할 수 있는  
Cost effective 한 Solution 필요



확장과 계산이 가능한  
예측 모델 필요

Deep  
Learning  
결과

## PayPal은 해결방안으로 Deep Learning 기법을 도입

- Low-level의 추상적 개념을 이해할 수 있도록 지원
- 기존과 다른 고도의 복잡성을 지닌 function에 대하여 학습할 수 있도록 지원
- 이미지, 영상 프로세싱, 실물 인식에 사용 가능

Deep  
Learning  
기법의  
효과

- 용이한 확장성
- 우월한 퍼포먼스
- 유연한 활용도
- 다른 Big Data Framework와 연동 가능
- 단순한 인터페이스

# DNN 활용 사례 – 국내 증권사 CASE



이탈 Data  
Featuring

이탈 유형에 따른 분석 모델링

유형별 원인분석 및  
활용

