AI, ML, 딥러닝

1. AI 그리고 ML

2016년 알파고의 등장으로 AI라는 연구가 핫 해졌습니다.

AI라는 것은 인간의 사고를 모방하기 위한 인간처럼 사고할 수 있게 연구하는 일종의 학문이라고도 말할 수 있는데요

AI의 기법 중 ML은 AI처럼 구현하기에 적합한 기법으로 알려지게 됩니다.

ML은 컴퓨터가 스스로 학습하는 즉, 기존의 Data를 이용해 앞으로의 일을 예측하는 기법입니다.

ML은 크게 패턴인식, 이상감지, 예측의 일을 하고 있습니다.

사람과 비교해봤을 때 사람이 사고하는 것과 유사하다는 점을 알게 되었습니다.

ML의 알고리즘 방식은 여러가지가 있는데요,

크게 지도학습과 비지도학습으로 나눌 수 있으며

지도 학습은 (예제를 주고 컴퓨터를 가르치는 학습 기법)  
 일반화 선형모델  
 의사결정트리  
 랜덤포레스트  
 점진적 부스트 머신  
 딥러닝

비지도 학습은 (사람의 개입이 없이 컴퓨터 스스로 데이터를 분류하는 기법)   
 클러스터링  
 비정상탐지  
 자원 축소

강화 학습

현재의 상태를 '강화' 하는 방법으로 프로그램이 구현되는 기법 입니다. S 라는 현재 상태에 A 라는 액션을 취했을 때 R 이라는 보상이 주어지고 프로그램은 이 R 을 최대 값으로 맞추는 방식으로 학습이 진행 됩니다.

등이 있습니다.

ML이 어떻게 동작하는지 잠깐 살펴보자면

예를 들어 바나나를 보여주면 ML이 학습을 합니다. 학습 시 “특징”을 찾아냅니다.

학습한 것을 모델로 만들어 두었다가 바나나를 보면 판단을 해서 “바나나”를 말하게 합니다.

기존의 규칙기반 인공지능에서는 특징을 알려주면 답을 말을 할 수 있지만,

ML은 특징을 스스로 찾아낸다는 의미입니다.

최소한의 모델을 주어야 함.

“선형 방정식의 형태?”

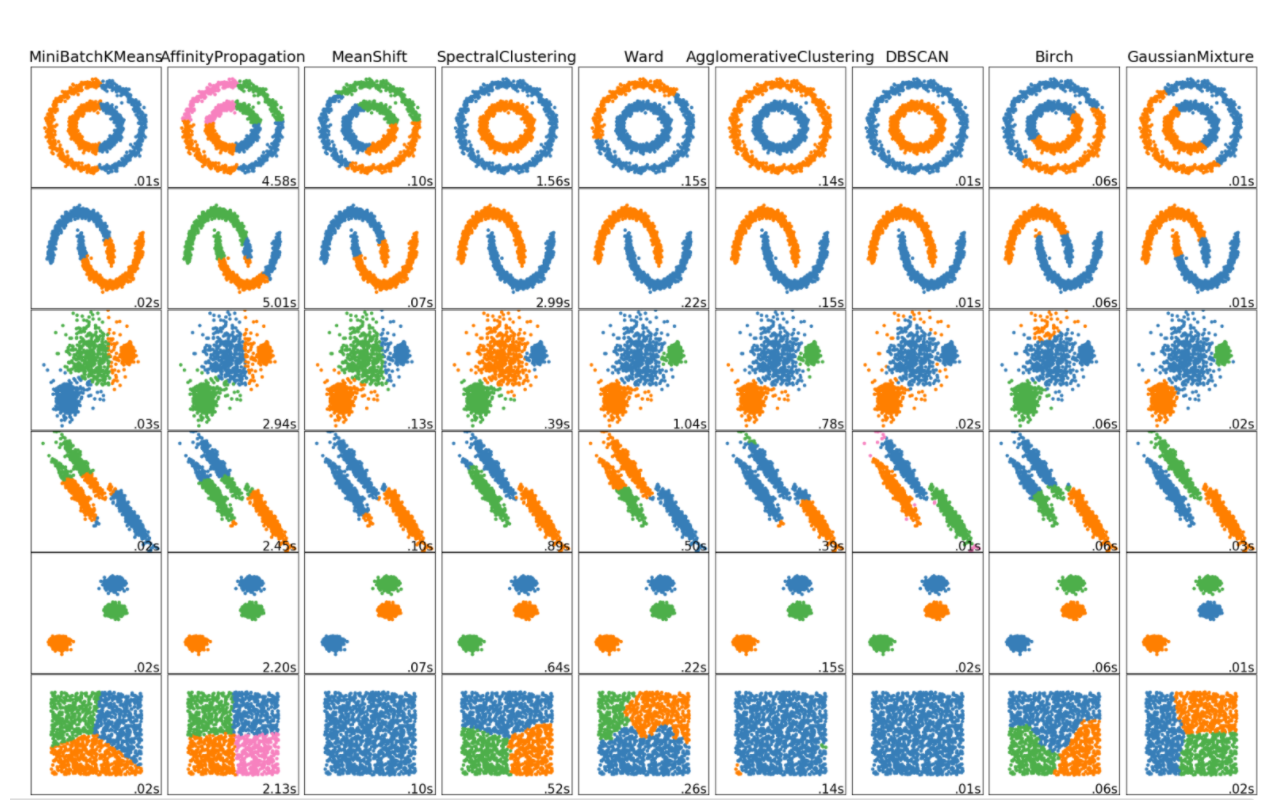
“인간의 뇌 구조를 딴 신경망의 형태”

데이터 / 모델 / 훈련 시간 => 모델 을 가지고 다른 곳에 쓰면 됨.

정리하자면 알고리즘을 통해 데이터를 분석하고 학습을 하여 판단을 하게 됩니다.

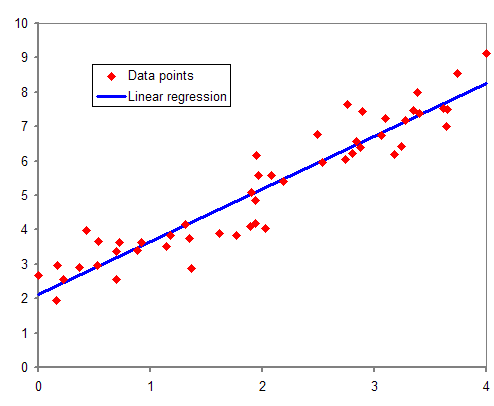


출처 : <https://jixta.wordpress.com/2015/07/17/machine-learning-algorithms-mindmap/>



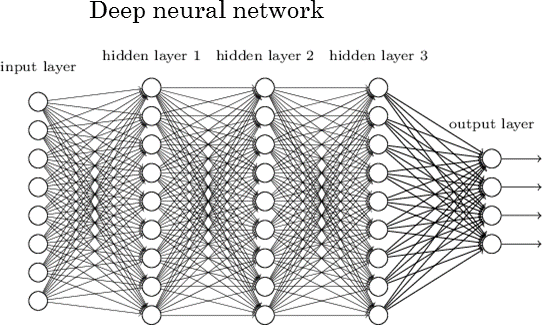
<https://bcho.tistory.com/1203>

클러스터링



선형 회귀

출처 : <https://ko.wikipedia.org/wiki/%EC%84%A0%ED%98%95_%ED%9A%8C%EA%B7%80#/media/File:Normdist_regression.png>



출처 : <https://cyberx.tistory.com/128>

딥러닝이란?

그렇다면 딥러닝은 무엇일까

딥러닝은 인간의 뇌의 동장방식을 착안하여 개발한 학습 방법으로 “인공신경망” 이라고 함.

딥러닝은 머신러닝 모델중 하나이다.

머신러닝의 모델 중 사람의 뇌를 본 뜬 구조가 있는데 인공신경망이라고도 하고 뉴럴 네트워크라고도 함.

굉장히 작은 처리하는 애들이 신경망처럼 연결되어 있어서 네트워크가 되어있는 형태

인공신경망 모델 구조를 써서 머신러닝 하는 걸 딥러닝이라고 함.

딥러닝은 1940년대부터 개발되었지만 90년대 부터는 한계가 보이기 시작했습니다.

인공신경망의 연구는 토론토 대학의 제프리 한톤교수가 전처리 과정을 해결하고 나서 부터 인정을 받았습니다.

왜?

1. GPU 병렬 컴퓨팅츼 등장,과 학습방법의 진보, CPU와 GPU가 병렬처리가 가능해 질만큼 기술의 발달 CuDA( 효율적으로 이용한 언어구조) 들의 개발,   
   과거 수 개월이 소요되었던 기계학습 과정은 GPU, 분산처리 환경의 발전 등**월등히 높아진 컴퓨팅 파워** 덕분에 이제 단지 수 시간, 혹은 몇 분만에 처리가 가능
2. 크게 3가지로 분류가 가능  
   1 ) 비지도 학습 방법 이용한 전처리 => 사과/ 고양이/ 사람 => 구분(클러스터링)  
     
   2 ) convolution 뉴럴 네트워크의 진화  
   ML은 데이터에서 특징을 뽑아내고 지식화 시키는 과정에서   
   특징을 뽑아내는 것이 기계학습의 성능을 좌우함.   
   CNN 기술이 발달  
     
   3 ) 시계열 Data를 위한 리커런트 뉴럴네트워크 – 시간의 흐름에 따라 변하는 DATA처리(주가, 움직임, 비디오)



출처 : <http://www.lgblog.co.kr/life-culture/business/116698>

딥러닝을 가장우월한가? NO 많은 양의 data와 많은 컴퓨팅 자원이 충족될 때 가장 효과적  
핸드폰이나 웨어러블 환경은 no

실시간 => 다른 방식이용.

결론 : AI와 ML과 DL에 대한 개념이 적립이 되었습니다.

ML안 모델을 공부하는 것이 목표이고 여기서의 핵심은

얼마나 효율성을 높일 수 있느냐와 얼마나 최적화 할 수 있느냐가 관건

예 ) 뭔가 핵심 주제에 대한 자연어 처리 업무를 적은 비용으로 쉽고 처리하느냐

ㅇㅖ) 우버에서 호출을 했을 때 예약 취소가 안날 활률을 계산 해서 얼마나 매칭을 최적화 가능한가.

딥러닝이 아닌 것도 많이 쓴다. 다른 모델의 장단점이 있다.

의사결정 트리, 서포트 벡터 머신 등등 다양한 방식이 있따.

그렇다면 왜 텐서플로우냐

1. 머신러닝 중 딥러닝 프로그램을 쉽게 구현
2. 자바, 고, 파이썬 지원 c++로 작성, 파이썬이 가장 편함
3. 커뮤티니 활발 (토치, 카페, 체이너.. 등 다양한 라이브러리 존재)
4. 클라우드 서비스, 깃 논문구현체도 가장 우선시 됨.

참조 1 : <https://www.youtube.com/watch?v=aF03asAmQbY&t=286s>

참조 2: <http://slownews.kr/41461>

참조 3: <https://kofboy2000.tistory.com/10>

참조 4, 5 : 모두의 딥러닝, 3분 딥러닝