**인공지능, 머신 러닝, 딥러닝의 차이점?**

인공지능 기술의 탄생

인공 지능이라는 개념은 1956년 미국 다트머스 대학에 있던 존 매카시 교수가 개최한 다트머스 회의에서 처음 등장했으며, 최근 몇 년 사이 폭발적으로 성장하고 있는 중이랍니다. 특히 2015년 이후 신속하고 강력한 병렬 처리 성능을 제공하는 GPU의 도입으로 더욱 가속화되고 있죠. 갈수록 폭발적으로 늘어나고 있는 저장 용량과 이미지, 텍스트, 매핑 데이터 등 모든 영역의 데이터가 범람하게 된 ‘빅데이터’ 시대의 도래도 이러한 성장세에 큰 영향을 미쳤습니다.

인공지능

머신러닝

딥러닝

인공지능

인공지능은 기계를 지능적으로 만드는 과학입니다.

기계는 문제를 해결할 때 알고리즘을 기반으로 문제를 해결하게 되는데, AI 알고리즘은 규칙이 생성되는 방식에서 기존 알고리즘과 차이가 있습니다.

기존 알고리즘은 개발자가 소프트웨어가 수신하는 각 유형의 입력값들에 대한 출력을 정의하는 특정 규칙을 설정하는 반면에 AI 알고리즘은 자체 규칙 시스템을 구축하게 됩니다.

이는 AI를 통해 컴퓨터가 사람에게 전적으로 의존했던 작업을 스스로 해결할 수 있음을 의미합니다.

머신러닝

인공지능의 하위 집한 개념인 머신러닝은 정확한 결정을 내리기 위해 제공된 데이터를 통하여 스스로 학습할 수 있습니다. 처리될 정보에 대해 더 많이 배울 수 있도록 많은 양의 데이터를 제공해야 합니다.

즉, 빅데이터를 통한 학습 방법으로 머신러닝을 이용할 수 있습니다. 머신 러닝은 기본적으로 알고리즘을 이용해 데이터를 분석하고, 분석을 통해 학습하며, 학습한 내용을 기반으로 판단이나 예측을 합니다. 따라서 궁극적으로는 의사 결정 기준에 대한 구체적인 지침을 소프트웨어에 직접 코딩해 넣는 것이 아닌, 대량의 데이터와 알고리즘을 통해 컴퓨터 그 자체를 ‘학습’시켜 작업 수행 방법을 익히는 것을 목표로 한답니다.

딥러닝

딥 러닝은 인공신경망에서 발전한 형태의 인공 지능으로, 뇌의 뉴런과 유사한 정보 입출력 계층을 활용해 데이터를 학습합니다. 그러나 기본적인 신경망조차 굉장한 양의 연산을 필요로 하는 탓에 딥 러닝의 상용화는 초기부터 난관에 부딪혔죠. 그럼에도 토론토대의 제프리 힌튼(Geoffrey Hinton) 교수 연구팀과 같은 일부 기관에서는 연구를 지속했고, 슈퍼컴퓨터를 기반으로 딥 러닝 개념을 증명하는 알고리즘을 병렬화하는 데 성공했습니다. 그리고 병렬 연산에 최적화된 GPU의 등장은 신경망의 연산 속도를 획기적으로 가속하며 진정한 딥 러닝 기반 인공 지능의 등장을 불러왔죠.

딥 러닝으로 훈련된 시스템의 이미지 인식 능력은 이미 인간을 앞서고 있습니다. 이 밖에도 딥 러닝의 영역에는 혈액의 암세포, MRI 스캔에서의 종양 식별 능력 등이 포함됩니다. 구글의 알파고는 바둑의 기초를 배우고, 자신과 같은 AI를 상대로 반복적으로 대국을 벌이는 과정에서 그 신경망을 더욱 강화해 나갔습니다.

머신러닝과 가장 큰 차이점은 딥러닝은 분류에 사용할 데이터를 스스로 학습할 수 있는 반면 머신 러닝은 학습 데이터를 수동으로 제공해야한다는점이 딥러닝과 머신러닝의 가장 큰 차이점입니다.

즉 딥러닝은 머신러닝을 달성하기 위한 방법이라고 볼 수 있다.

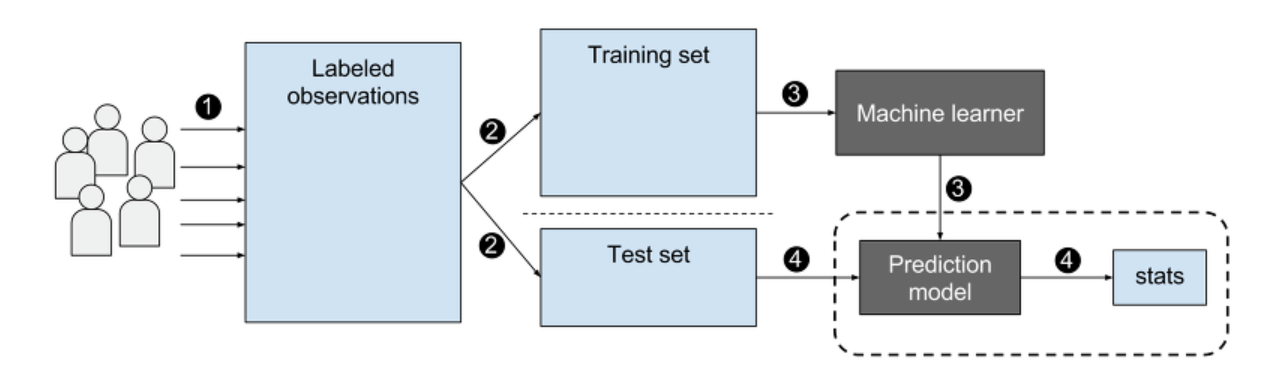
Reference

* <https://wendys.tistory.com/136?category=765608>

**머신러닝의 종류?**

머신러닝에는 세가지 종류가 있다.

지도학습(Supervised Learning), 비지도학습(Unsupervised Learning), 강화학습(Reinforce Learning)



1. 지도학습(Supervised Learining)

가장 기본이 되고, 가장 구현하기 쉬운 알고리즘이 지도학습으로 일단 컴퓨터에게 문제(Feature)와 정답(Label)이 있는 데이터(Training Set)를 학습 시킨 후, 운영 데이터(Test Set)를 분류하거나 맞추는 것이 비지도 학습이다.

예를 들어서, 어떤 지역의 원룸 매매 가격을 맞추는 지도학습을 만들어보고자 한다. 그래서 현재 거래가 완료된 데이터를 학습을 시켜서, 현재 등장한 매물을 보고 정가에 팔릴 것인지를 예측하는 시스템을 만들고자 한다.

이때 **매매가 이미 완료된 데이터**가 학습에 사용되는 **학습 데이터(Training Set)**이다. 이 학습 데이터는 Feature와 Label이 있어야 하는데 Feature는 지역, 층수, 엘레베이터유무, 평수, 방수, 화장실 개수, 건축년도 등등 상당히 많은 수가 있을 것이고, Label은 매매가격이다.

1. 비지도학습(Unsupervised Learning)

비지도 학습은 비지도 한다는 말처럼 사람의 개입이 없이 컴퓨터가 알아서 어떠한 결과를 도출해 내는 것을 뜻한다. 지도 학습은 우리가 결과를 안다는 가정하에 컴퓨터가 그에 맞춰서 문제를 해결하는 것을 원한다고 한다면 비지도는 우리도 뭔지 모르는 무언가를 컴퓨터가 해결하길 바라는 것이다.

예를 들어, 어떤 로그 파일에 대한 데이터 분석을 수행했는데 아무런 Insight를 못찾았다고 했는데 컴퓨터의 비지도 학습 알고리즘을 돌려보니 사람들이 인지하지 못했던 어떤 분류를 했다던지, 문제를 풀어버리는 행위이다.

비지도 학습에 대표적인 방법이 클러스터링(Clustering) 즉 군집화가 존재한다. 너무도 많은 Feature로 인해서 인간이 어떤식으로 분류를 해야 할지 모르는 상태에서 Clustering 알고리즘을 활용하여 컴퓨터에게 맡기면 새로운 방식으로 인간이 알지 못하는 방식의 분류를 수행할 수 있다.

물론 클러스터링이 언제나 정답은 아니다. 간혹 왜 이렇게 분류를 했지? 아무리 찾아도 모르는 경우가 있고 그럴 때는 다시 어느정도의 Feature들을 컴퓨터가 이해하기 쉽게 선작업을 해야 하는 경우가 있다.

비지도 학습은 인간이 인지하지 못하는 부분에 많이 쓰이기 때문에 금융권에서 FDS 시스템(금융 위험 이상 탐지 시스템)에서도 많이 쓰이고, 해커들이 침입했을 때의 패턴을 스스로 분석하여 해커들의 침입을 인지할 수도 있다. 즉 지도 학습을 사용하기 힘든 부분에 컴퓨터의 머리를 빌려 학습을 해 나가는 것이 비지도 학습이라고 쉽게 생각하면 된다

출처: <https://needjarvis.tistory.com/195> [자비스가 필요해]

1. 강화학습(Reinfoce Learining)

사실 강화 학습은 상당히 애매한 영역이다. 2가지로 분류해야 한다면 지도학습의 카테고리에 들어가야 하는 기법인데 최근 들어서 알파고의 학습으로 유명해진 기법이다.

강화 학습의 핵심은 바로 보상(Reward)이다. 컴퓨터는 이 보상을 받는 행위를 위해서 스스로 문제점을 찾아나가게 되는데 초반에는 인간의 어느정도의 개입이 들어갈 수 있다.

지도학습과 비지도학습은 컴퓨터가 스스로 성장하는 모습이 쉽지 않는 모양이지만, 강화학습은 목표가 보상의 최대치가 되는 행동을 하자라는 것이다. 예를 들어, 게임을 한다면 최종스코어(Max Score)를 최대치를 만드는 것이 임무일 수 있고, 이 최대치를 만들기 위해서 끊임없이 노력을 하게 된다.

알파고 같은 경우는 초반에 바둑 프로들이 둔 기보를 통해서 성장을 하다가 나중에 알파고 대 알파고의 대결로 새로운 기보를 만들어 내고, 끊임없이 성장해 나가게 되는데 이 이유는 지도 학습이 아닌 강화 학습이기 때문에 가능한 것이다.

출처: <https://needjarvis.tistory.com/195> [자비스가 필요해]  
Reference

* <https://needjarvis.tistory.com/195>

**Docker란**

reference

초보자를 위한 도커 안내 블로그 : <https://subicura.com/2017/01/19/docker-guide-for-beginners-1.html>

**pytorch 기반의 딥러닝 기초**

reference

코랩 사용법(김태영님 블로그) :  <https://tykimos.github.io/2019/01/22/colab_getting_started/>

코랩 사용법 강의 : <https://youtu.be/inN8seMm7UI>