

# 제1장: 딥러닝이란 무엇인가

**Soyoung Park**

Pusan National University  
Department of Statistics

**`soyoung@pusan.ac.kr`**

# Assessment and homework

## 과제와 평가

- 논문발표 20%
- Weekly assignments 30%
- Final projects 50%

## Text books

- 매주 제공되는 강의노트, R 실습 파일
- 텐서플로 케라스를 이용한 딥러닝 (박유성 지음)

# 인공지능, 머신러닝, 딥러닝



<sup>1</sup><https://samstory.coolschool.co.kr/zone/story/modi/streams/76601>

# 인공지능, 머신러닝, 딥러닝

## 인공지능

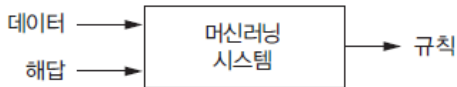
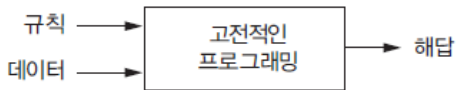
- 1950년대 출현 → 컴퓨터를 '생각'하게 만들수 있을지에 대한 노력으로 시작
- 
- 머신러닝, 딥러닝을 포괄하며,

# 인공지능, 머신러닝, 딥러닝

## 머신러닝을 시작하게한 질문들:

- '수행방법을 아는 모든 일을 넘어서, 지정한 작업을 수행하는 방법을 계산기가 스스로 학습할 수 있는가?'
- 프로그래머가 직접 데이터 처리 규칙을 작성하는 대신, 컴퓨터가 데이터를 보고

# 인공지능, 머신러닝, 딥러닝



## 보기 1.2 머신러닝: 새로운 프로그래밍 패러다임

# 인공지능, 머신러닝, 딥러닝

## 머신러닝

- 머신러닝 시스템은 명시적으로 프로그래밍되는것이 아니라 →
- 과업과 관련된 많은 examples를 접한 후, 이 속에서
- 예시:
- 고속 처리 하드웨어와 대규모 dataset의 활용이 가능해진 1990년대 이르러서야 인기를 끌기 시작

# 인공지능, 머신러닝, 딥러닝

## Machine learning vs. Statistical learning

- 머신러닝은
- 복잡한 dataset
- 머신러닝, 특히 딥러닝은 상대적으로
- 머신러닝 분야에서는



# Machine learning vs. Statistical learning

## Machine learning:

- Machine learning is the field that deals with
- Programs and systems can accomplish tasks
- For example, image recognition technology often relies on machine learning algorithms that parse huge numbers of pictures, learning to identify objects and other features within those images over time and after analyzing large volumes of image data.

# Machine learning vs. Statistical learning

## Statistical learning:

- Statistical learning focuses on
- Since statistical learning may be used to develop the underlying models that govern how a machine learning algorithm understands data,
- One basic example of this in action is a linear regression algorithm, which is a type of machine learning algorithm that was developed based on the principles of statistics.

---

<sup>1</sup><https://onlinestats.canr.udel.edu/machine-learning-vs-statistics/>

# Machine learning vs. Statistics

*“The major difference between machine learning and statistics is their purpose. Machine learning models are designed to make the most accurate predictions possible. Statistical models are designed for inference about the relationships between variables.”*

---

<sup>1</sup><https://towardsdatascience.com/the-actual-difference-between-statistics-and-machine-learning-64b49f07ea3>

# 인공지능, 머신러닝, 딥러닝

## 머신러닝 필요규칙

- 입력 데이터 점:
- 예상 출력 사례:
- 알고리즘의 성공적인 수행 여부를 측정하는 방법:

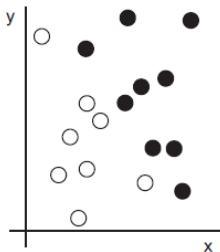
# 머신러닝

## 입력 데이터 점

- 머신러닝 모델은
- 이는
- 머신러닝과 딥러닝의 핵심과제는
- 즉, 예상하는 결과와 더 가깝게 접근할 수 있도록
- 표현(representation)이란 데이터를 대표(represent)하거나, 데이터를 부호화(encode)하기 위해 데이터를 색다르게 살펴보는 방법
- ex) 컬러 이미지는 RGB, HSV형식으로 부호화 가능

# 인공지능, 머신러닝, 딥러닝

점의 좌표  $(x, y)$ 를 취해 그 점이 검은색인지, 흰색인지를 출력할 수 있는 알고리즘을 개발하려고 한다고 가정해보자.



보기 1.3 일부 표본 데이터

# 인공지능, 머신러닝, 딥러닝

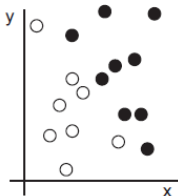
## 머신러닝 필요규칙

- 입력은
- 기대하는 결과는
- 알고리즘이 올바른 작업을 수행하는지 여부를 측정하는 방법은

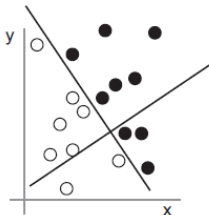
# 인공지능, 머신러닝, 딥러닝

흰색점과 검은색 점을 깔끔하게 구분하는 데이터의 새로운 표현 →

1: 원자료

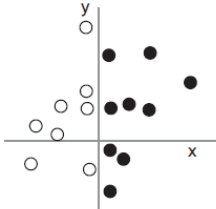


2: 좌표 변환



보기 1.4 좌표 변환

3: 더 나은 표현



점의 색에 관한 규칙:



# 인공지능, 머신러닝, 딥러닝

## 머신러닝

- 좌표변환의 방식을
- 머신러닝에서 학습이란
- 모든 머신러닝 알고리즘은 데이터를 주어진 작업을 수행하기에 더 나은 표현으로 바꾸는 변환 방식을

# 인공지능, 머신러닝, 딥러닝

## 딥러닝(deep learning)

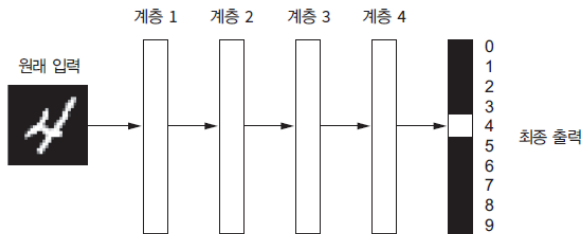
- 딥러닝은 머신러닝의 특정 하위 분야
- 표현을 더 의미있게 만들어가는
- 딥(deep)이라는 문구는 심층적인 이해수준을 언급하는것이 아님
- 딥(deep)은
- layered representations learning, hierarchical representations learning

# 인공지능, 머신러닝, 딥러닝

## 딥러닝과 통계학

- 통계적 머신러닝은
- 딥러닝은
- 공통점:
- 딥러닝은
- 딥러닝은

# 인공지능, 머신러닝, 딥러닝



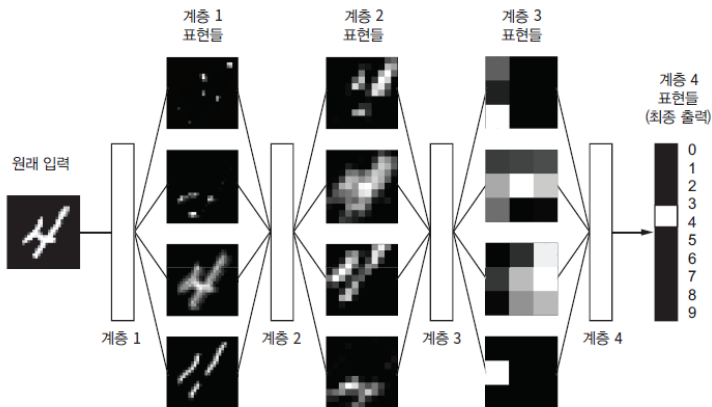
보기 1.5 숫자 분류를 위한 심층 신경망

# 인공지능, 머신러닝, 딥러닝

## 딥러닝(deep learning)

- 현대 딥러닝에서는
- 딥러닝에서
- 신경망이라는 생물학적 단어를 차용하였지만, 딥러닝이 뇌모델은 아님

# 인공지능, 머신러닝, 딥러닝



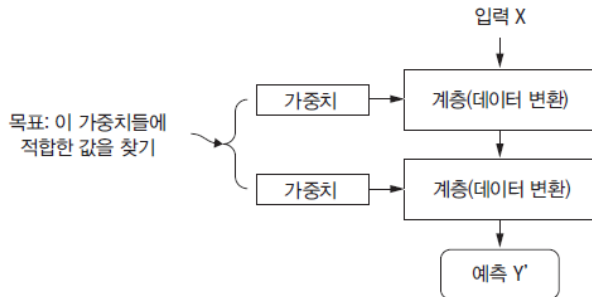
보기 1.6 숫자 분류 모델이 학습한 심층 표현들

# 인공지능, 머신러닝, 딥러닝

## 딥러닝(deep learning)

- network는 숫자 이미지를
- 심층망(deep network)은 다단식 정보 증류 연산(information distillation operation)으로 생각할 수 있음, 즉 정보가 연속적인 증류기를 거치면서 점차 순수해진다고 해석
- 즉 데이터 표현을
- 따라서 딥러닝에 관한 과한 믿음이나 오해는 버리고 딥러닝은

# 인공지능, 머신러닝, 딥러닝



보기 1.7 신경망의 가중치들에 의해 신경망이 파라미터화된다.

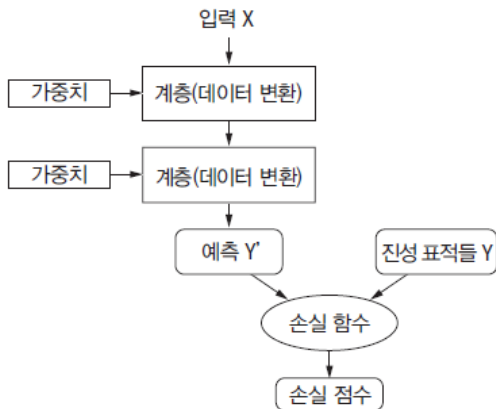


# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식

- 어떤 계층에 입력되는 데이터는 해당 계층에 의해 특정화 (specification)된 다음,
- 가중치:
- 즉, 계층에 의해 구현된 변환이
- 따라서 딥러닝에서의 학습이란,

# 인공지능, 머신러닝, 딥러닝



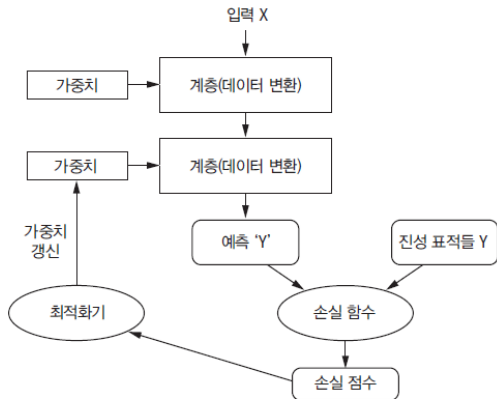
보기 1.8 손실 함수는 망 출력의 품질을 측정한다.

# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식

- 신경망 출력을 제어하려면,
- 이것을
- 손실 함수는

# 인공지능, 머신러닝, 딥러닝



보기 1.9 손실 점수는 가중치를 조정하는 피드백 신호로 사용된다.

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 기능적인 측면에서의 이해

- 
- 
-

# 인공지능, 머신러닝, 딥러닝

딥러닝의 개념적인 측면에서의 이해 - 데이터의 구조에 따라

- MLP (multilayer perception) -
- CNN (convolutional neural networks) -
- RNN (recurrent neural networks) -

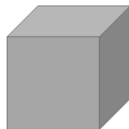
# 인공지능, 머신러닝, 딥러닝



vector



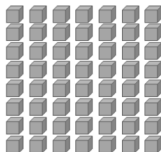
matrix



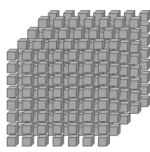
3d-tensor



4d-tensor



5d-tensor



6d-tensor

<sup>1</sup><https://rekt77.tistory.com/102>

# 인공지능, 머신러닝, 딥러닝

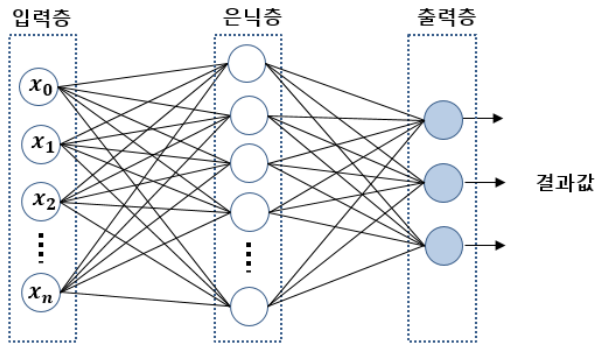
## 텐서(tensor)

- 텐서:
- 0차원 텐서:
- 1차원 텐서(1D 텐서):
- 2차원 텐서(2D 텐서):
- 3차원 텐서(3D 텐서):
- 4차원 텐서(4D 텐서):



# 딥러닝 작동 방식

- 딥러닝은
- 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)



<sup>1</sup><https://m.blog.naver.com/samsjang/221030487369>

# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식-입력층(input layer)

- 이 텐서들이 은닉층에 전달되어 목적변수를 잘 설명하도록

# 인공지능, 머신러닝, 딥러닝

딥러닝 작동 방식-은닉층(hidden layer)

- 특성변수가 은닉층에 전달되면,
- 
-

# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식-출력층(output layer)

- 입력된 매개변수를
- 출력층에 입력된 매개변수를
- ex)

# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식-모수 최적화

- 예측된 목적변수값은 실제 목적변수 값과 비교하여 일종의 거리인
- 
- 
- 역전파는

# 인공지능, 머신러닝, 딥러닝

## 딥러닝 작동 방식-과대적합(overfitting)

- 학습데이터(training data) -
- 검증데이터(validation data) -
- 시험데이터(testing data) -

# 인공지능, 머신러닝, 딥러닝

## 딥러닝(머신러닝)과 통계학의 가장 큰 차이점 -일반화(generalization)

- 통계학에서의 일반화:
- 딥러닝(머신러닝)에서의 일반화:
- 따라서
-

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 업적

- 딥러닝은 개발된지 비교적 오래되었지만, 2010년 초반에 이르러서야 눈에 띄게 됐다.
- 최근 10년간 딥러닝은 혁신이나 다를 바 없는 성과를 거두었는데, 특히 보는일이나 듣는일과 같은 지각문제
- 인간에게는 자연스럽고 직관적인 것으로 보이지만, 오랫동안 기계로는 파악이 힘들었던 기술문제등에서 놀라운 결과를 나타냄



# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 업적

- 이미지를 사람이 분류하는 정도에 가깝게 분류
- 음성을 사람이 알아듣는 수준에 가깝게 인식
- 필기체를 사람이 인식하는 수준에 가깝게 인식
- 더 나은 기계 번역
- 텍스트를 더 자연스러운 음성으로 변환
- 구글 나우, 아마존 알렉사와 같은 디지털 비서 .. 등

딥러닝으로 무엇까지 할 수 있는지는 지금도 연구 대상! →

# 인공지능, 머신러닝, 딥러닝

## 머신러닝의 역사..

- 확률론적 모델링: naive Bayes algorithm, logistic regression (분류 알고리즘)
- 초기 신경망: Yann LeCun이 convolutional neural network과 역전파 (1980년대)를 결합
- 커널 메서드(Kernel method): Support vector machine (분류 알고리즘)
- SVM: 대형 데이터셋까지 확장해 적용하기 어렵고, 이미지 분류와 같은 문제에 적용 불가능

# 인공지능, 머신러닝, 딥러닝

## Decision trees, Random forest, Gradient boosting machine

- Decision tree: 입력데이터 점을 분류하는데 사용, 결정 트리를 시각화하여 해석하기 좋음
- Random forest: 다수의 전문화된 DT를 구축하고, 그 결과를 종합한다
- → 모든 shallow machine learning 작업용으로 두번째로 우수.
- Gradient boosting machine: 일반적으로 약한 예측 모형을 종합하는 방식을 기반으로 하며, 기울기 증폭(gradient boosting)을 사용함
- → 기울기 증폭 기법이 결정 트리에 적용되면, 속성이 유사하면서도 거의 항상 RF보다 성능이 훨씬 더 뛰어난 모델이 된다. (Kaggle 대회에서 가장 흔히 사용되는 알고리즘!)

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 차별점



# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 차별점: 특징공학의 자동화

- Shallow learning (RF, SVM 등)에서는 입력 데이터를 보통 고차원 비선형 투영법 또는 단순한 변환 등으로 변환이 필요했다. 즉,
- 반면, 딥러닝은 특징공학의 단계를
  - 
  - 
  - 
  -

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 차별점

- 딥러닝 모형은
- 딥러닝이 혁신적인 점은
- 
- 합동특징학습(joint feature learning)을 통해, 내부 특징 중 하나를 조정할 때마다 자동으로 의존하는 다른 모든 특징이 사람의 개입 없이 자동으로 변경됨. 모든 특징이 단일 피드백 신호에 의해 지도됨.

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 두가지 필수 특성

- 표현이 여러 계층을 거치는 동안에
- 중간과정에서 나타나게 되는 여러 표현을 신경망의 중간에 자리 잡고 있는 여러 계층들이

# 인공지능, 머신러닝, 딥러닝

## 왜 딥러닝은 지금인가?

- 하드웨어 (빨라진 CPU, GPU의 개발)
- Dataset 및 벤치마크 (인터넷 등장으로 인한 초대규모 데이터셋 수집 및 배포 가능)
- 알고리즘의 발전 (기울기 전파-gradient propagation의 개발)



# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 민주화

- 도구의 민주화: 파이썬, R 등 기본적인 스크립트 기술을 가지고도 고급 딥러닝 연구 수행 가능
- Theano, TensorFlow, keras, XGboost 등 사용자 친화적 라이브러리의 등장
- 특히 2015년 개발된 Keras는 수많은 대학원생, 연구원, 기업들이 현장으로 진출하는데 솔루션 제공

# 인공지능, 머신러닝, 딥러닝

## 딥러닝의 밝은 미래의 이유:

- 단순성(simplicity): 복잡하고 부서지기 쉬운 공학 수준의 파이프라인을 end-to-end 학습 모델로 대체함으로써 특징공학을 불필요하게 만듦
- 확장성(scalability): 딥러닝은 GPU, TPU의 병렬처리에 매우 적합하므로, 소규모 배치들을 반복하는 식으로 training하기 때문에, 대규모 데이터셋도 처리 가능
- 다기능성 및 재사용성(versatility and resuability): 딥러닝은 다른 머신러닝 기법과 달리, 밑바닥부터 재시작하지 않고도 추가 데이터를 훈련할 수 있음.
- 예를들어, 이미지 분류로 training한 딥러닝 알고리즘을 영상 처리 작업 공정 한곳에 끼워넣을 수 있음.