

[Deep Learning] 딥러닝의 기초 - 1.2 딥러닝 이전: 머신러닝의 간략한 역사 — 나무늘보의 개발 블로그

노트북: 첫 번째 노트북

만든 날짜: 2021-01-25 오후 9:55

URL: <https://continuous-development.tistory.com/225>

---

Deep Learning

## [Deep Learning] 딥러닝의 기초 - 1.2 딥러닝 이전: 머신러닝의 간략한 역사

2021. 1. 21. 02:18 수정 삭제 공개

# 1.2 딥러닝 이전: 머신러닝의 간략한 역사



### 1.2.1 확률적 모델링

확률적 모델인은 통계학 이론을 데이터 분석에 응용한 것이다.

가장 잘 알려진 알고리즘 중 하나는 나이브 베이즈 알고리즘이다.

나이브 베이즈는 입력 데이터의 특성이 모두 독립적이고 가정하고 베이즈 정리를 적용하는 머신 러닝 분류 알고리즘이다.

### 1.2.2 초창기 신경망

딥러닝의 기원

신경망의 핵심 아이디어는 1950년대에 시작됐으나 대규모 신경망을 훈련시킬 수 있는 효과적인 방법을 오랜 기간 동안 찾지 못했다.

1980년 중반에 역전파 알고리즘을 재발견하고 신경망에 이를 적용하기 시작했다. 이 알고리즘은 경사 하강법 최적화를 사용하여 연쇄적으로 변수가 연결된 연산을 훈련하는 방법이다.

첫 번째 신경망 애플리케이션은 1989년 벨 연구소에서 나왔다. 이때 얀르 쿤은 초창기 합성곱 신경망과 연전파를 연결하여 손글씨 숫자 이미지를 분류하는 문제에 적용했다.

LeNet이라 부르는 이 신경망은 우편 봉투의 우편 번호 코드를 자동으로 읽기 위해 1990년대 미국 우편 서비스에 사용 되었다.

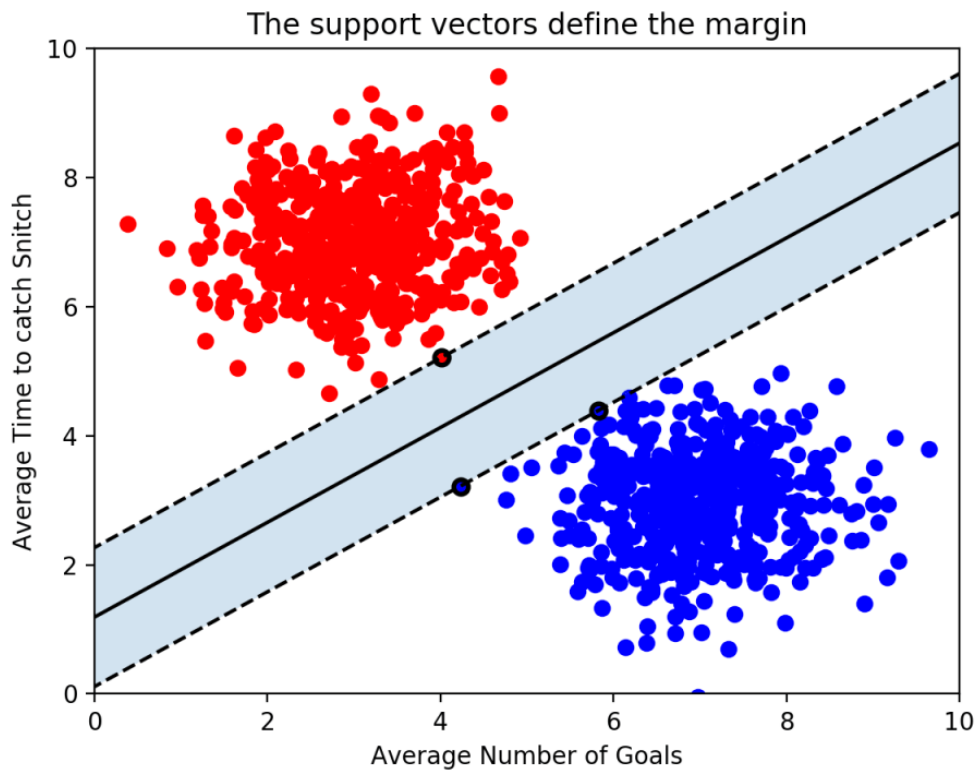
### 1.2.3 커널 방법

1990년대에 신경망은 어느정도 관심 얻었지만, 머신 러닝의 새로운 접근 방법인 커널 방법이 인기를 얻자 신경망은 빠르게 잊혀졌다.

커널 방법은 분류 알고리즘의 한 종류를 말하며 그중 서포트 벡터 머신이 가장유명하다.

SVM은 분류 문제를 해결하기 위해 2개의 다른 범주에 속한 데이터 포인트 그룹 사이에 좋은 결정 경계를 찾는다.

결정 경계는 훈련 데이터를 2개의 범주에 대응하는 영역으로 나누는 직선이나 표면으로 생각할 수 있다.



SVM이 결정 경계를 찾는 과정은 두 단계이다.

- 1.결정 경계가 하나의 초평면(hyperplane)으로 표현될 수 있는 새로운 고차원 표현으로 데이터를 매핑한다.
- 2.초평면과 각 클래스의 가장 가까운 데이터 포인트 사이의 거리가 최대가 되는 최선의 결정 경계를 찾습니다. 이 단계를 마진 최대화라고 부름

다. 이렇게 함으로써 결정 경계가 훈련 데이터셋 이외의 새로운 샘플에 잘 일반화되도록 도와준다.

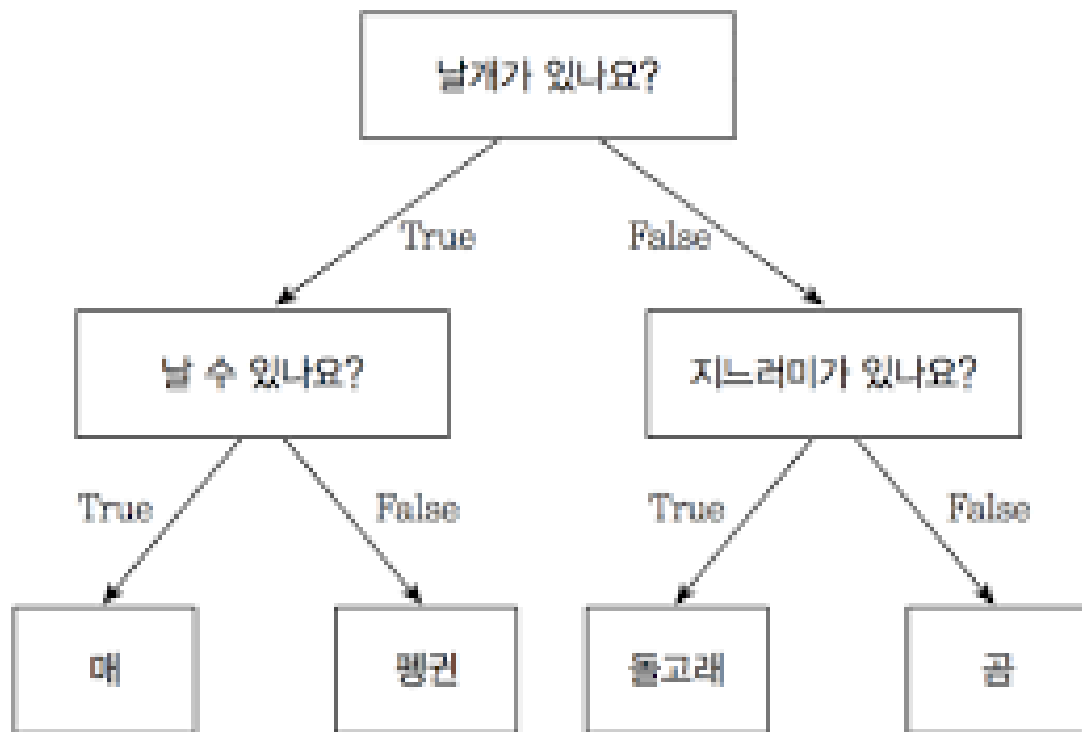
여기서 우리는 좋은 초평면을 찾기 위해 포인트의 좌표를 구할 필요 없이 두 데이터 포인트 사이의 거리를 계산할 수만 있으면 된다.

커널함수를 사용하면 이를 효율적으로 계산할 수 있다.

커널 함수는 원본 공간에 있는 두 데이터 포인트를 명시적으로 새로운 표현으로 변환하지 않고 타깃 표현 공간에 위치했을 때의 거리를 매핑해 주는 계산 가능한 연산이다.

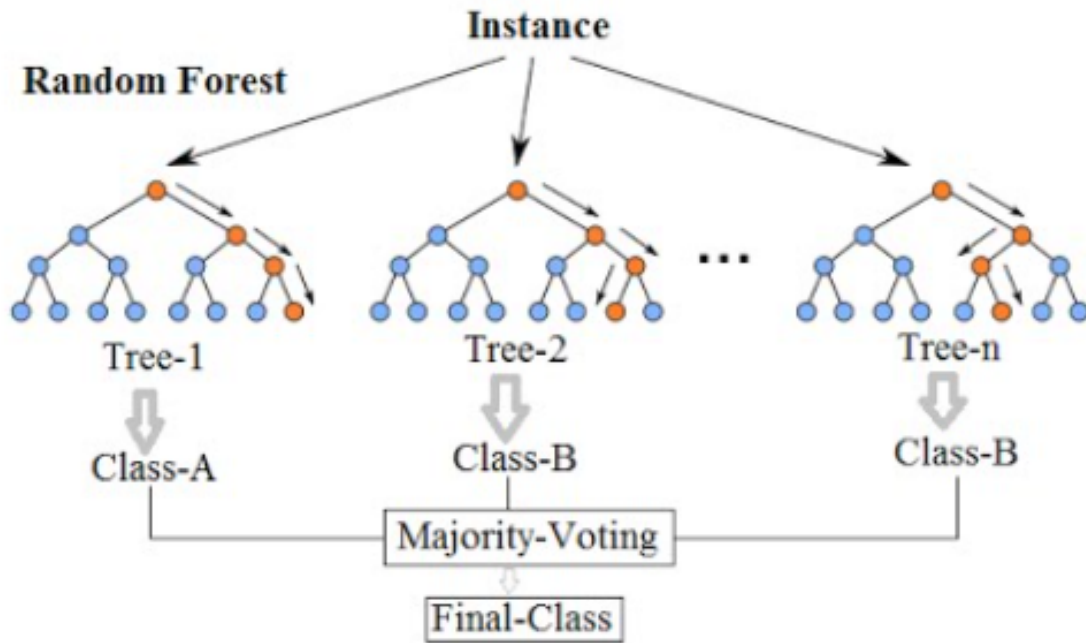
SVM은 간단한 분류문제에는 최고 수준의 성능을 보였지만 대용량 데이터셋에 확장되기 어려웠고 시각에 관련된 문제에 좋은 성능을 내지 못했다.

## 1.2.4 결정 트리, 랜덤 포레스트, 그래디언트 부스팅 머신

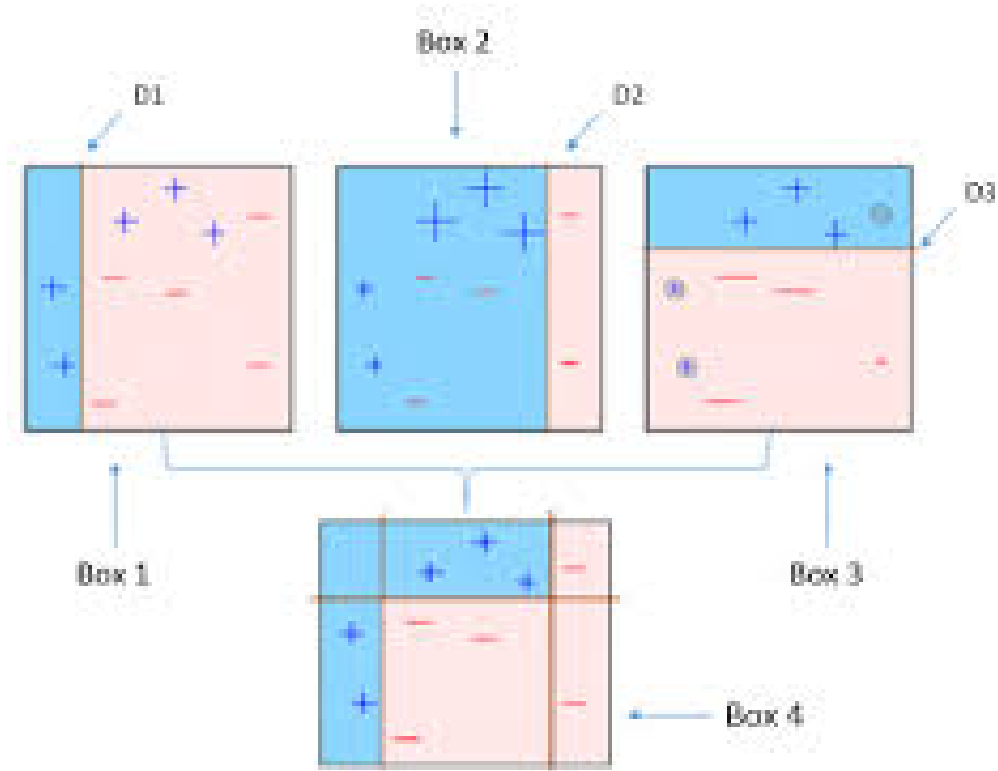


**결정트리**는 flow차트 같은 구조를 가지며 입력 데이터 포인트를 분류하거나 주어진 입력에 대해 출력 값을 예측한다.

# Random Forest Simplified



**랜덤포레스트 알고리즘**은 결정트리 학습에 기초한 것으로 안정적이고 실전에 유용하다. 서로 다른 결정트리를 많이 만들고 그 출력을 앙상블하는 방법을 사용한다.



**그래디언트 부스팅 머신**은 약한 예측모델인 결정트리를 앙상블하는 것을 기반으로 하는 머신 러닝 기법이다. 이 알고리즘은 이전 모델에서 놓친 데이터 포인트를 보완하는 새로운 모델을 반복적으로 훈련함으로써 머신러닝 모델을 향상 시킨다.

## 1.2.5 다시 신경망으로

2011년 IDSIA의 댄 크리슨이 GPU로 훈련된 심층 신경망으로 학술 이미지 분류 대회에서 우승한 것을 시작으로 다시 뜨기 시작했다.

이때부터 매년 ImageNet 대회는 심층 합성곱 신경망이 우승을 차지했다. 심층 합성곱 신경망이 모든 컴퓨터 비전 작업의 주력 알고리즘이 되었다.

## 1.2.6 딥러닝의 특징

딥러닝이 이렇게 빠르게 확산된 주된 이유는 많은 문제에서 더 좋은 성능을 내고 있기 때문이다.

그것과 더불어 딥러닝은 머신러닝에서 가장 중요한 단계인 특성 공학을 완전히 자동화하기 때문에 문제를 더 해결하기 쉽게 만들어준다.

머신러닝에서 복잡한 문제에 필요한 잘 정제된 표현은 일반적으로 얻기 어렵다. 그래서 사람이 초기 입력 데이터를 여러 방식으로 변환해야 한다.

즉 데이터의 좋은 표현을 수동으로 만들어야 한다.

이를 특성공학이라고 한다.

특성공학에 대해 더 자세히 말하자면 머신러닝에서 분석을 위해 데이터에 대한 도메인 지식을 활용하여 특성을 만들어내는 과정이다.

이것을 Feature Engineering이라고 한다.

머신러닝에서 사용되는 특성공학 기법이다.

- 차원 축소(Dimension Reduction)
  - 협의적으로는 아래서 설명하는 특성 추출을 차원 축소와 동일한 개념으로 이해
  - 광의적으로는 아래서 설명하는 두가지 기법을 아우르는 상위 개념
  - 특성 선정(Feature Selection)
    - 특성 랭킹(Feature Ranking) 또는 특성 중요도(Feature Importance)에 따라 선정
    - 분석가의 사전 배경 지식을 이용하거나, 랜덤 선정 및 성능 측정을 반복하는 방법 사용



- 특성 추출(Feature Extraction)
  - 특성을 단순히 선택하는 것이 아니라 특성들의 조합으로 새로운 특성을 생성
  - PCA는 특성 추출 기법에 속함
- 스케일링(Scaling)
  - 변수의 분포가 편향되어 있을 경우, 변수 간의 관계가 잘 드러나지 않는 경우 범위 변환
  - 전체 값에 로그나 제곱루트 적용하는 방법 등
- 변형(Transform)
  - 기존 존재하는 변수의 성질을 이용해 다른 변수를 만드는 방법
  - 날짜별 판매 데이터에 주말 여부를 추가하는 등, 분석가에 경험에 따라 추가 또는 변형
- 비닝(Binning)
  - 연속형 변수를 범주형 변수로 변환
  - 33살, 47살 -> 30대, 40대
- 더미(Dummy)
  - 범주형 변수를 연속형 변수로 변환 가능
  - 기존 특성이나 특성의 조합을 규칙에 따라 수치화하여 Dummy Data로 추가

딥러닝을 사용하면 특성을 직접 찾는 대신 한 번에 모든 특성을 학습할 수 있다. 머신러닝 작업 흐름을 매우 단순화시켜 주므로 고도의 다단계 작업 과정을 하나의 간단한 엔드-투-엔드 딥러닝 모델로 대체할 수 있다.

3개의 층을 가진 모델에서 최적의 첫 번째 표현 층은 하나의 층이나 2개의 층을 가진 모델에서 최적의 첫 번째 층과는 달라야한다. 딥러닝의 변환 능력은 모델이 모든 표현 층을 순차적이 아니라 동시에 공동으로 학습하

게 만든다. 이런 공동 특성 학습 능력 덕택에 모델이 내부 특성 하나에 맞추어질 때마다 이에 의존하는 다른 모든 특성이 사람이 개입하지 않아도 자동으로 변화에 적응하게 된다.

모든 학습은 하나의 피드백 신호에 의해 시작된다. 즉 모델의 모든 변화는 최종 목표를 따라가게 된다. 이 방식은 모델을 많은 중간영역으로 나누어 복잡하고 추상화된 표현을 학습시킬 수 있기 때문에 얇은 학습모델을 탐욕적으로 쌓은 것보다 강력하다.

딥러닝이 데이터로부터 학습하는 방법에는 두 가지 중요한 특징이 있다. 층을 거치면서 점진적으로 더 복잡한 표현이 만들어진다는 것과 이런 점진적인 중간 표현이 공동으로 학습된다는 사실이다.

각 층은 상위 층과 하위 층의 표현이 변함에 따라서 함께 바뀐다. 이 특징이 머신러닝접근보다 딥러닝이 훨씬 성공하게 된 이유이다.

## 1.2.7 머신러닝의 최근 동향

구조적인 데이터인 경우에는 그래디언트 부스팅이 사용되고, 딥러닝은 이미지 분류 같은 시각에 관한 문제에 사용된다.

본 내용은 케라스 창시자에게 배우는 딥러닝 책을 참고한 내용입니다

---

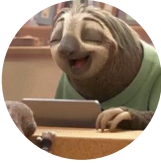
**'Deep Learning' 카테고리의 다른 글**

[Deep Learning] 딥러닝의 기초 - 1. 3 왜 딥러닝일까?

[Deep Learning] 딥러닝의 기초 - 1.2 딥러닝 이전: 머신러닝의 간략한 역사

딥러닝의 역사

딥러닝의 특징



나아무늘보

혼자 끄적끄적하는 블로그 입니다.