

# 머신러닝

---

I. 머신러닝 개요 / 응용

II. 지도 학습

# 인공지능과 머신러닝

---

- ◆ 인공지능은 지능형 에이전트의 설계와 분석에 관한 작업
- ◆ 지능적인 행동을 위해서는 지식(예: 환경에 대한 모델)이 필요함
- ◆ 주어진 작업을 수행하는데 필요한 지식을 명확히 기술하는 것은 어려운 일이며 불가능한 경우도 존재
- ◆ 머신러닝(ML) : 컴퓨터가 다음 작업을 통해 지식 습득하는 과정
  - 데이터 관측
  - 데이터를 활용하여 모델 구축
  - 모델을 세계에 대한 가설, 또는 문제 해결 소프트웨어의 일부로 사용

# 학습 에이전트

---

- ♣ 학습을 통해 에이전트의 결정 메커니즘을 수정하고 성능을 개선
  - 개선 대상 구성 요소는?
  - 에이전트가 가지고 있는 사전 지식 중 모델에 영향을 미치는 지식은?
  - 사용할 수 있는 데이터는 어떤 것이며 그 데이터에 대한 가용한 피드백은?
- ♣ 환경은 시간이 지남에 따라 변화하며, 학습은 이러한 변화에 적응해야 함
- ♣ 환경이 알려지지 않은 상황에서는 학습이 필수적임

# 머신러닝 응용분야

---

- Agriculture
- Anatomy
- Adaptive websites
- Affective computing
- Banking
- Bioinformatics
- Brain-machine interfaces
- Cheminformatics
- Citizen science
- Natural language processing
- Natural language understanding
- Online advertising
- Optimization
- Recommender systems
- Robot locomotion
- Search engines
- Sentiment analysis
- Computer networks
- Computer vision
- Credit-card fraud detection
- Data quality
- DNA sequence classification
- Economics
- Financial market analysis<sup>[75]</sup>
- General game playing
- Handwriting recognition
- Sequence mining
- Software engineering
- Speech recognition
- Structural health monitoring
- Syntactic pattern recognition
- Telecommunication
- Theorem proving
- Time series forecasting
- User behavior analytics
- Information retrieval
- Insurance
- Internet fraud detection
- Linguistics
- Machine learning control
- Machine perception
- Machine translation
- Marketing
- Medical diagnosis

# 머신러닝 모델 종류

---

## ♣ 지도 학습

- ◆ Naïve Bayes classifier
- ◆ Nearest neighbor methods
- ◆ Linear models
- ◆ Decision trees
- ◆ Neural networks
- ◆ Support vector machines
- ◆ Ensemble learning

## ♣ 확률론적 그래프 모델

- ◆ Bayesian networks
- ◆ Markov random fields

## ♣ 비지도 학습

- ◆ Clustering: mixture models, K-means, hierarchical clustering
- ◆ Principal component analysis
- ◆ Independent component analysis

## ♣ 순차적 데이터 학습

- ◆ HMMs
- ◆ Recurrent neural networks

## ♣ 마코프 결정 과정

## ♣ 강화학습

# 지도 학습

에이전트는 입력-출력 쌍을 관찰하고 입력에서 출력으로 매핑하는 함수를 학습

**문제**  $N$  개의 쌍으로 구성된 **학습 데이터 셋**이 아래와 같이 주어짐:

$$(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$$

각 데이터 쌍이 미지의 함수  $y = f(x)$ 에 의해 생성되었다고 할 때, 원 함수  $f$ 를 근사하는 함수  $h$ 를 찾아라.

가능한 함수들로 구성된 **가설 공간**  $\mathcal{H}$ 로부터  
추론된 **가설** (또는 **모델**)

데이터 생성에 관한  
사전 지식에 따라 선택

$y_1, y_2, \dots, y_n$ : 모델이 예측할 **ground truth**

# 최적(Best-Fit) 함수

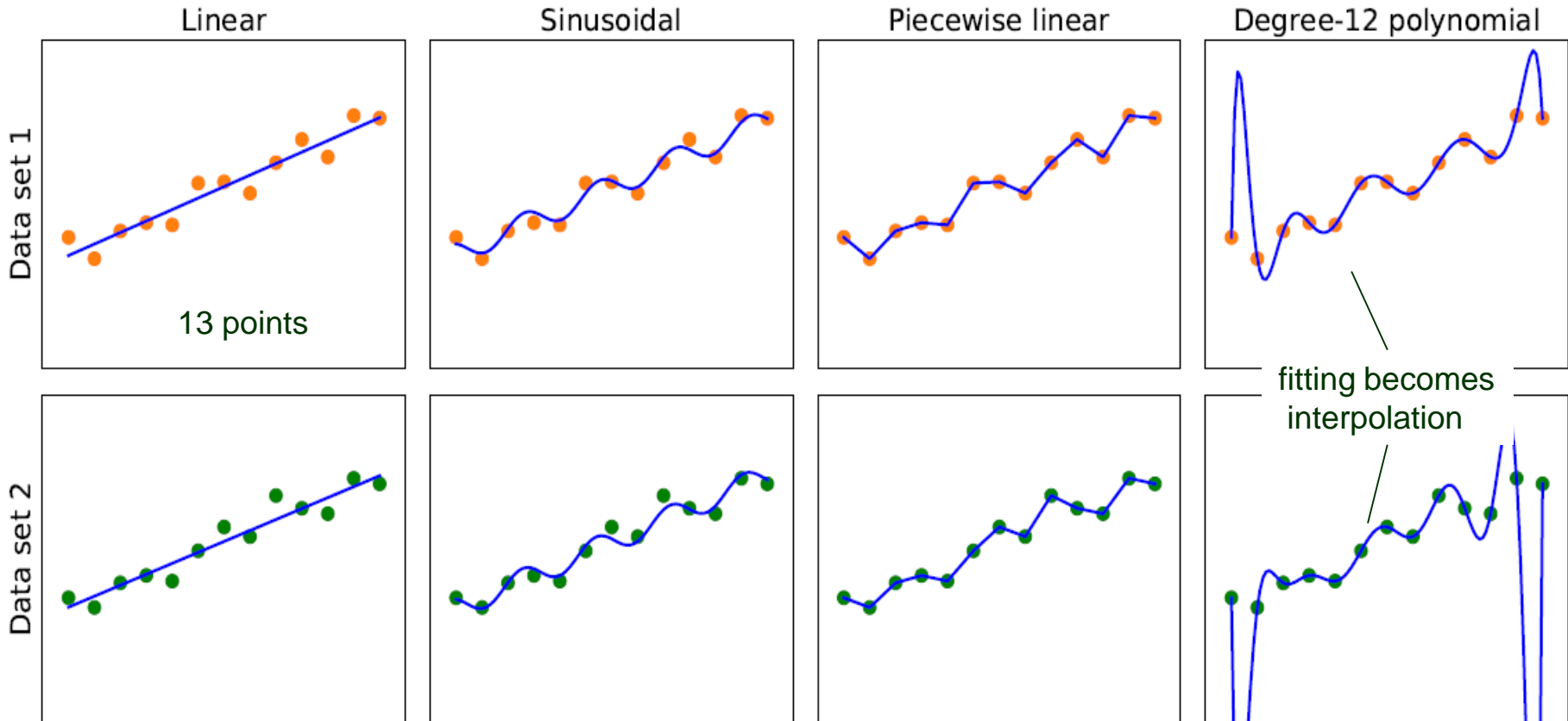
---

- ♣ 데이터에 잡음이 포함되어 있기 때문에 ground truth와 정확히 일치하는 예측 함수는 구할 수 없음, 즉  $h(x_i) = y_i$ , for  $1 \leq i \leq N$ 는 존재하지 않음
- ♣ 대신, 각  $h(x_i)$  가  $y_i$ 와 근사적으로 가까운 최적 함수  $h$ 를 구함
- ♣  $h$ 의 성능을 제대로 측정하는 방법은 미지의 입력에 대한 평가하는 것

*테스트 세트*:  $(x_i, y_i)$  쌍의 두 번째 샘플

- ♣ 만약  $h$ 가 테스트 세트에 대해 높은 정확도를 유지한다면,  $h$ 는 잘 일반화되었다고 할 수 있음

# 적합화(Fitting)



$$h(x) = w_1x + w_0$$

$$h(x) = w_1x + \sin(w_0x)$$

$$h_i(x) = w_{i1}x + w_{i0},$$

$$1 \leq i \leq 12 \text{ such that}$$

$$h_j(x_{j+1}) = h_{j+1}(x_{j+1})$$

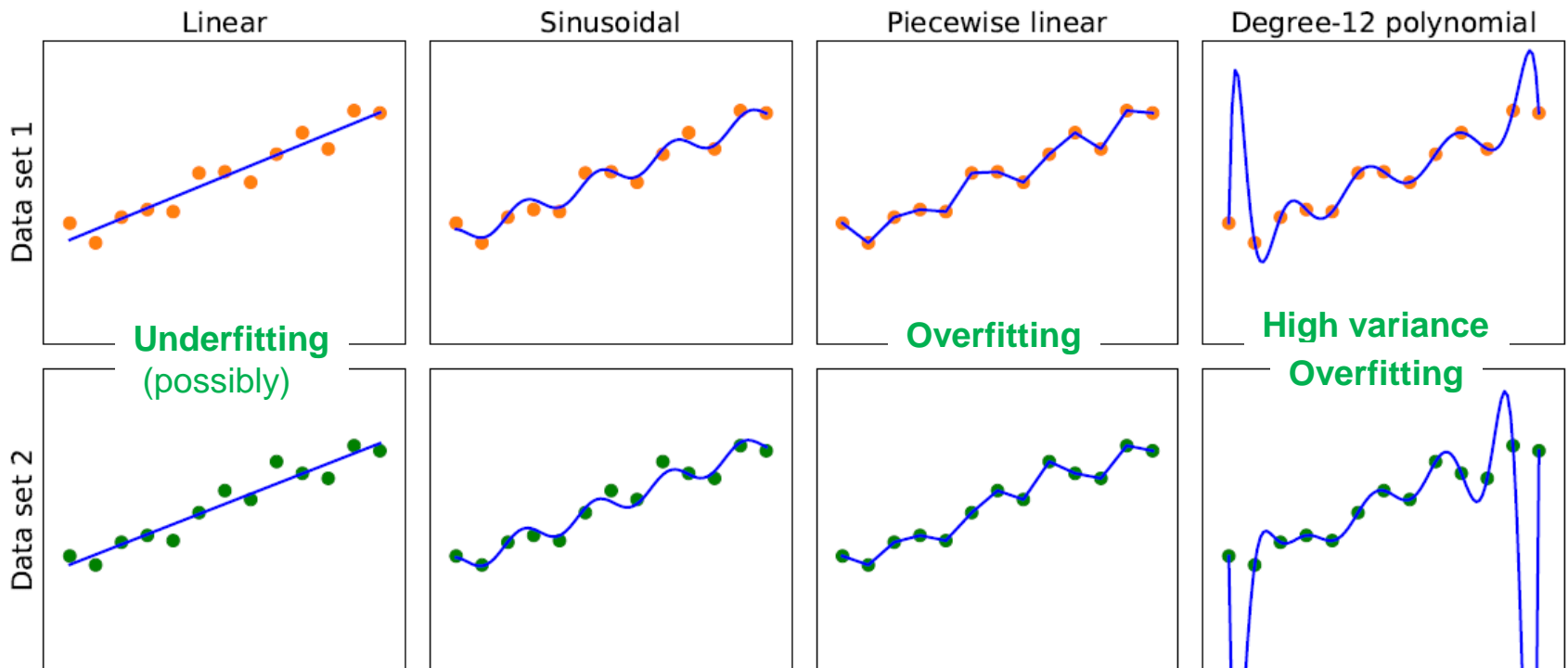
$$\text{for } 1 \leq j \leq 11.$$

$$h(x) = \sum_{i=0}^{12} w_i x^i$$



# 과소적합, 분산, 과적합

- **과소적합(underfitting)**: 가설이 데이터에서 패턴을 찾지 못한 경우
  - ◆ 이런 가설은 높은 편향과 낮은 분산을 가짐
- **분산**은 훈련 데이터가 바뀔때 따라 가설이 얼마나 변하는지를 나타냄
- **과적합(overfitting)**: 가설이 특정 훈련 세트에 너무 많은 주의를 기울이는 경우
  - ◆ 그런 가설은 낮은 편향과 높은 분산을 가짐



# 최적 가설

---

데이터에 의해 표현된 것과 데이터로부터 기대하는 것이 무엇인지에 따라 달라짐

주어진 데이터를 가장 잘 표현하는 가설을 선택:

$$h^* = \operatorname{argmax}_{h \in \mathcal{H}} P(h \mid data)$$

$$= \operatorname{argmax}_{h \in \mathcal{H}} P(data \mid h)P(h) \quad (\text{Bayes' rule})$$

가설 공간  $\mathcal{H}$  중 간단한 것이 더 선호됨

♠  $\mathcal{H}$ 의 표현 가능성이 더 풍부할수록 해당 가설 공간에서 좋은 가설을 찾는 계산 비용이 더 높아짐

♣ 학습 완료 후  $h$ 는 보통 평가 목적으로 사용됨

# 식당에서 자리 대기 문제

---

**Problem** 식당에서 빈 좌석이 나올 때 까지 기다릴지 결정하기

**Output:** 부울 변수 *WillWait* (빈 좌석 기다린다면 true).

**Input:** 이산적인 값을 갖는 열 개의 속성 벡터

1. *Alternate*: 주변에 적합한 다른 식당이 있는지
2. *Bar*: 식당에 편안한 바 공간이 있는지
3. *Fri/Sat*: 금요일이거나 토요일이면 true
4. *Hungry*: 지금 당장 배가 몹시 고프지
5. *Patrons*: 식당에 손님이 많은지 (values: *None*, *Some*, *Full*)
6. *Price*: 식당의 음식 가격대가 비싼지 (\$, \$\$, \$\$\$).
7. *Raining*: 밖에 비가 오는지
8. *Reservation*: 예약을 해 두었는지
9. *Type*: 식당 종류 (*French*, *Italian*, *Thai*, *burger*).
10. *WaitEstimate*: 호스트 예상 대기시간: 0-10분, 10-30분, 30-60분, 60분 이상

$2^6 \times 3^2 \times 4^2 = 9,216$  가지의 속성 값 조합이 존재

# 학습 예제

Example	Input Attributes										Output
	<i>Alt</i>	<i>Bar</i>	<i>Fri</i>	<i>Hun</i>	<i>Pat</i>	<i>Price</i>	<i>Rain</i>	<i>Res</i>	<i>Type</i>	<i>Est</i>	<i>WillWait</i>
<b>x<sub>1</sub></b>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Some</i>	<i>\$\$\$</i>	<i>No</i>	<i>Yes</i>	<i>French</i>	<i>0-10</i>	<i>y<sub>1</sub> = Yes</i>
<b>x<sub>2</sub></b>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Full</i>	<i>\$</i>	<i>No</i>	<i>No</i>	<i>Thai</i>	<i>30-60</i>	<i>y<sub>2</sub> = No</i>
<b>x<sub>3</sub></b>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Some</i>	<i>\$</i>	<i>No</i>	<i>No</i>	<i>Burger</i>	<i>0-10</i>	<i>y<sub>3</sub> = Yes</i>
<b>x<sub>4</sub></b>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Full</i>	<i>\$</i>	<i>Yes</i>	<i>No</i>	<i>Thai</i>	<i>10-30</i>	<i>y<sub>4</sub> = Yes</i>
<b>x<sub>5</sub></b>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Full</i>	<i>\$\$\$</i>	<i>No</i>	<i>Yes</i>	<i>French</i>	<i>&gt;60</i>	<i>y<sub>5</sub> = No</i>
<b>x<sub>6</sub></b>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>	<i>Some</i>	<i>\$\$</i>	<i>Yes</i>	<i>Yes</i>	<i>Italian</i>	<i>0-10</i>	<i>y<sub>6</sub> = Yes</i>
<b>x<sub>7</sub></b>	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>No</i>	<i>None</i>	<i>\$</i>	<i>Yes</i>	<i>No</i>	<i>Burger</i>	<i>0-10</i>	<i>y<sub>7</sub> = No</i>
<b>x<sub>8</sub></b>	<i>No</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Some</i>	<i>\$\$</i>	<i>Yes</i>	<i>Yes</i>	<i>Thai</i>	<i>0-10</i>	<i>y<sub>8</sub> = Yes</i>
<b>x<sub>9</sub></b>	<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>No</i>	<i>Full</i>	<i>\$</i>	<i>Yes</i>	<i>No</i>	<i>Burger</i>	<i>&gt;60</i>	<i>y<sub>9</sub> = No</i>
<b>x<sub>10</sub></b>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Full</i>	<i>\$\$\$</i>	<i>No</i>	<i>Yes</i>	<i>Italian</i>	<i>10-30</i>	<i>y<sub>10</sub> = No</i>
<b>x<sub>11</sub></b>	<i>No</i>	<i>No</i>	<i>No</i>	<i>No</i>	<i>None</i>	<i>\$</i>	<i>No</i>	<i>No</i>	<i>Thai</i>	<i>0-10</i>	<i>y<sub>11</sub> = No</i>
<b>x<sub>12</sub></b>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Full</i>	<i>\$</i>	<i>No</i>	<i>No</i>	<i>Burger</i>	<i>30-60</i>	<i>y<sub>12</sub> = Yes</i>

- ♣ 9,216가지 가능한 예제 중 단 12개의 실 사례만 주어짐
- ♣ 나머지 9,204개의 사례는 최선의 추측으로 채워야 함