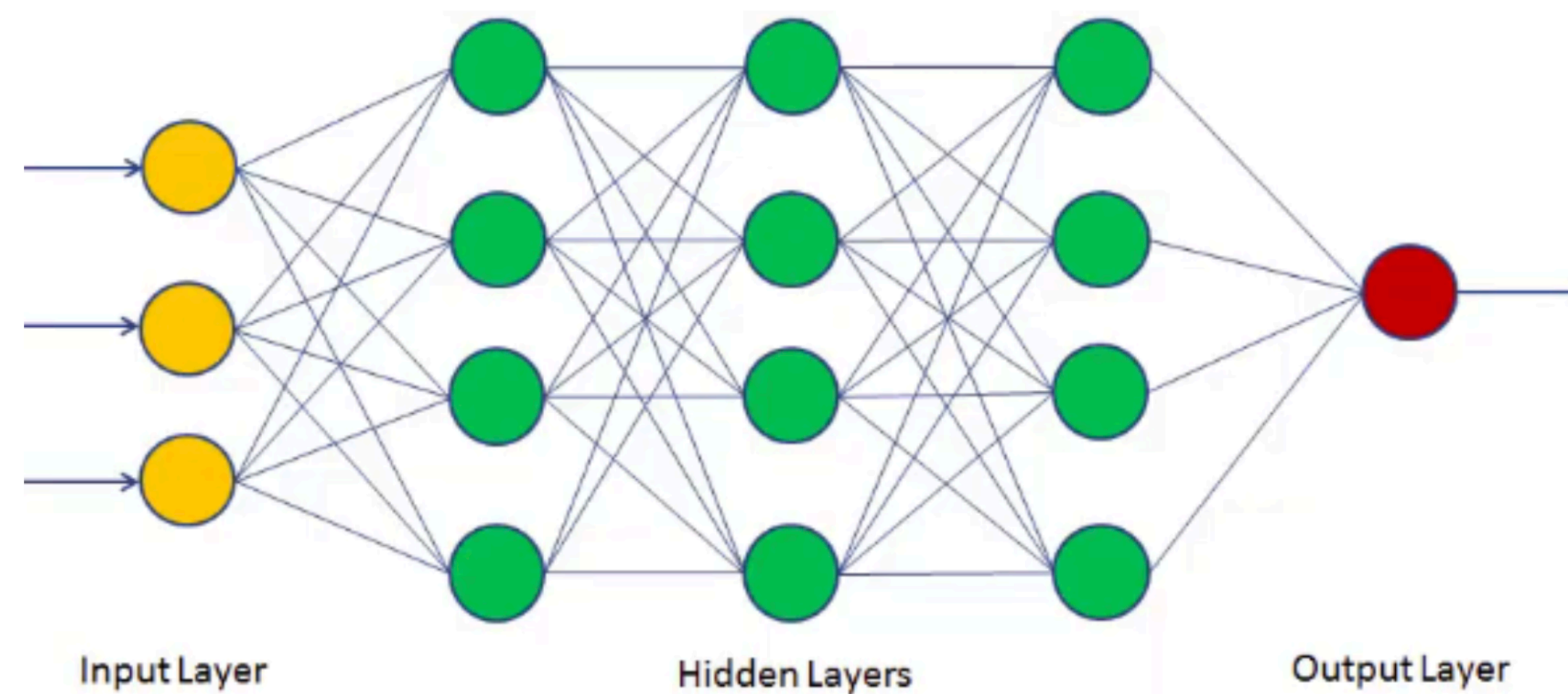
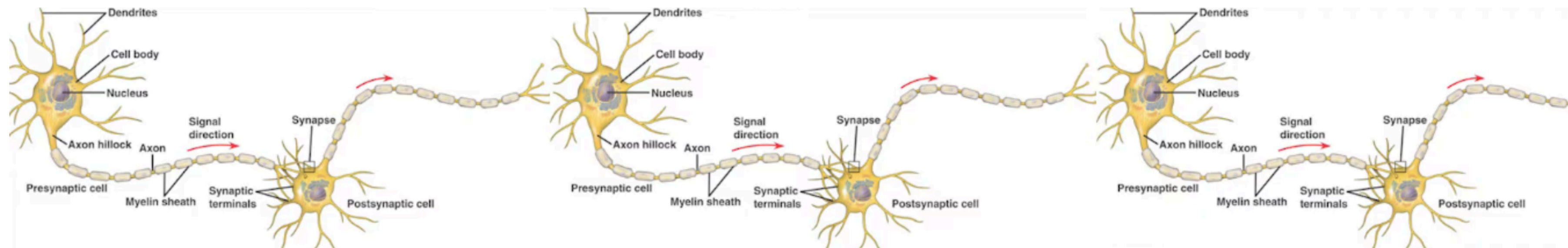


Deep Learning

- 머신러닝의 한 종류
- 여러 층을 가진 신경망(Neural Network)을 사용해 머신러닝을 수행 하는것
- 인공 신경망은 인간의 뇌신경망의 신호 체계를 본따서 만든 구조



Deep Learning과 Machine Learning의 차이점

- 가장 큰 차이점은 특징량(Features) 추출
- 머신러닝 방법은 사람이 이미지 내부에 있는 특징을 하나하나 지정하는 방식.
- 딥러닝에서는 학습데이터에서 기계가 자동으로 특징을 추출.
- 딥러닝은 스스로 학습을 하기 때문에 머신러닝 보다 많은 Data가 필요.
- 예를 들어 사과와 포도를 판별하는 경우 사람이 직접 색에 주목해서 살펴보라고 지정 했다면 딥러닝을 사용해 대량의 데이터를 학습 시키면 이러한 특징을 기계가 직접 찾아 냅니다. 이때 특징이 색이 될수도 있고 모양이 될수 도 있습니다.

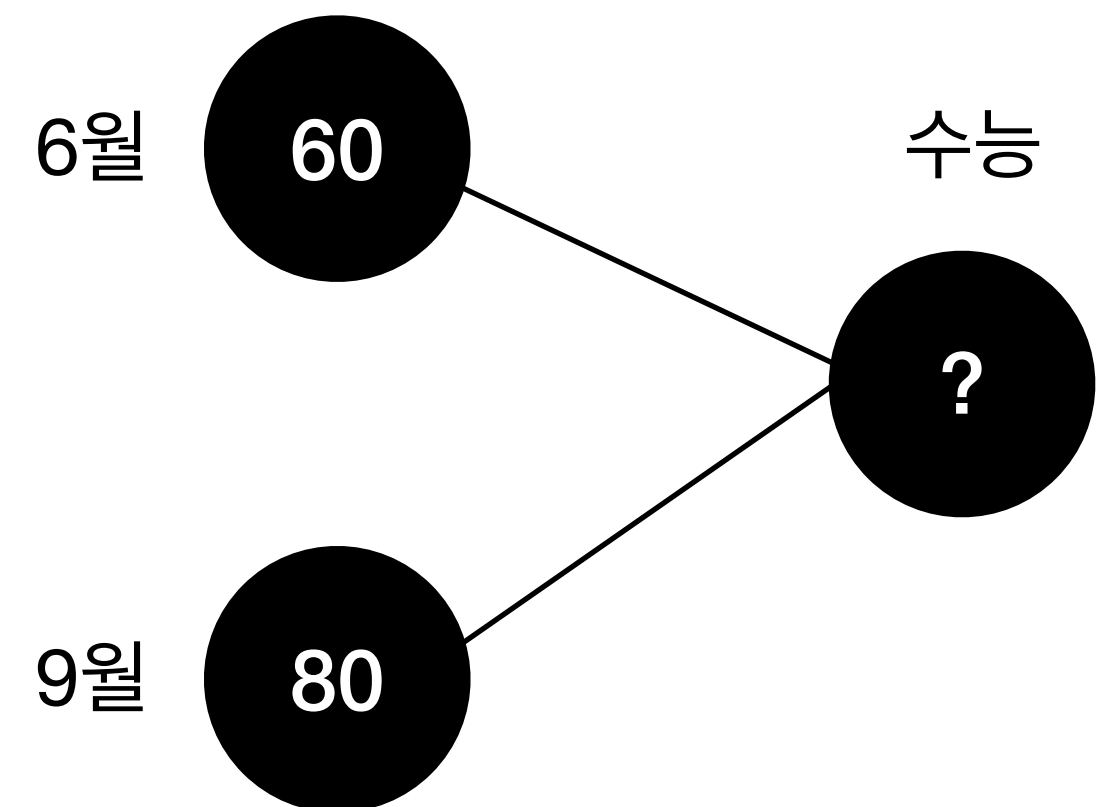
신경망(Neural Network)

- 입력층, 은닉층, 출력층을 지나 처리가 일어나고 최종적인 결과로 출력
- 각 층에서는 입력은 여러군데에서 받지만 출력은 하나의 유닛으로만 출력을 보냅니다.
- 이러한 신경망을 3개이상 중첩하면 Deep Neural Network(DNN)이라고 하며 이를 활용한 기계학습을 딥러닝 이라 한다.
- 딥러닝은 대량의 데이터를 입력해서 학습시킵니다.
- 학습이란 이러한 유닛(뉴런)의 접속 가중치를 조정하는 것을 의미

딥러닝이 머신러닝보다 우수한 분야

- | | | | |
|--|---|--------|-----|
| - Image classification
- Image Object Detection | ➡ | 자율 주행차 | CNN |
| - Sequence Data | ➡ | 번역기 | RNN |

Machine Learning 알고리즘



$(60 * 0.5) + (80 * 0.5)$

$(60 * 0.5) + (80 * 0.5) + \text{bias}(\text{절편, 편향 : 난이도, 가산점})$

$(6\text{월점수} * w1) + (9\text{월점수} * w2) + b$

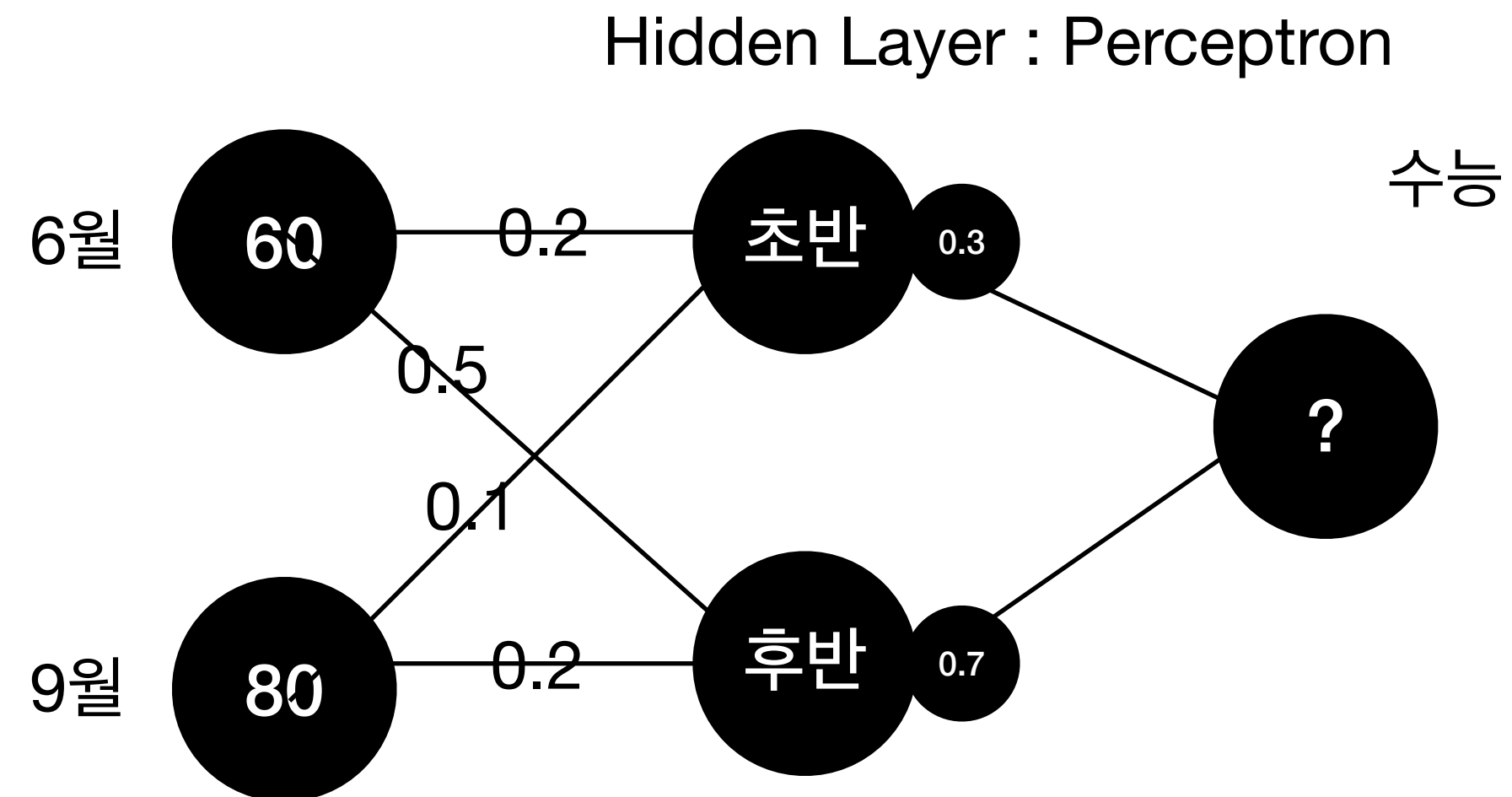
	6월	9월	수능	예측	오차
A	60	80	90	70	20
B	50	60	70	55	15
C	60	60	80	60	20

이 오차의 최소값을 찾아 w1, w2, b를 선정한다.

총 오차값을 계산하기 위한 알고리즘

- Cross-Entropy
- Binary cross entropy

Deep Learning 알고리즘



Deep Learning 을 이용한 계산

- 실제 Data를 이용하여 오차를 구한후 그 오차들의 합을 구한다.
- 이때 오차들의 합이 가장 적었을때의 w값을 취한다.
- 오차를 구할때 그냥 뺄셈만 이용하면 마이너스 값이 발생하는 경우가 발생한다.
- 이를 해결하기 위해 절대값을 이용하던가 평균제곱오차(오차값들의 제곱의 합계)를 계산한다.

Loss Function(손실함수)

- 평균제곱오차를 구하여 최소화 시키는 방법
- 모델의 정확도를 평가하는 함수

결과값이 정수를 구하고자 할때

$$\frac{1}{n} \sum (\hat{y} - y)^2$$

0~1과 같이 확률적인 값을 구하고자 할때

$$-\frac{1}{n} \sum [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

Activation Function(활성함수)

- Sigmoid 함수
- tanh(Hyperbolic tangent)
- ReLu(Rectified Linear Unit)

최적화(Optimization)

- 대표적인 방법은 경사 하강법
- 반복적으로 손실함수에 대한 모델 매개변수의 기울기(미분값)를 구한후 그 미분값의 반대 방향으로 매개 변수를 조절해 나가면 결국에는 최저 손실함수값에 도달한다.

학습률(Learning Rate)

- 학습률을 크게 설정하면 매개 변수 변경치가 커져서 최적의 값이 도달하지 않을수 도 있다.

Adagrad

- 각 매개변수에 각기 다른 학습률을 적용하고 변화가 많은 가중치는 학습률을 적게 설정하고, 변화가 적은 가중치는 학습률을 높게 설정하는 옵티마이저
- 자연어 처리에 적합

Adam

- Adagrad의 학습률 자율 조정과 효율적인 매개변수 변경 알고리즘을 조합한 알고리즘

딥러닝의 과대 적합

Drop Out

- 매개변수중 일정량을 학습 중간마다 무작위로 사용하지 않는 방법
- 앙상블 효과를 준다.

Early Stopping

- 반복 횟수를 높이면 학습시간이 길어져서 학습데이터만 좋은 현상이 발생
- 조기종료는 학습 횟수에 따라 검증 정확도가 꾸준히 떨어지는 시점이 발견되면 그 즉시 학습을 중단하고 최고점을 사용한다는 개념.