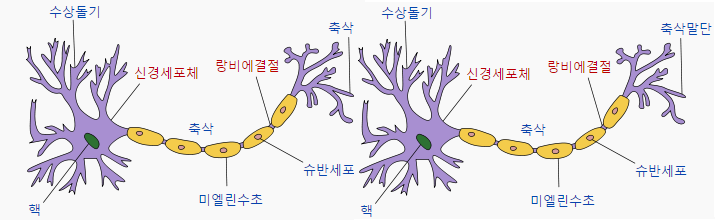
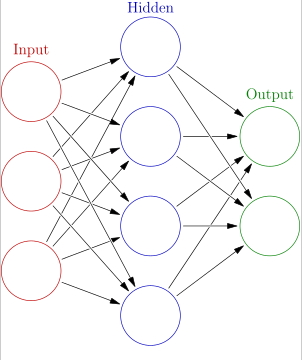
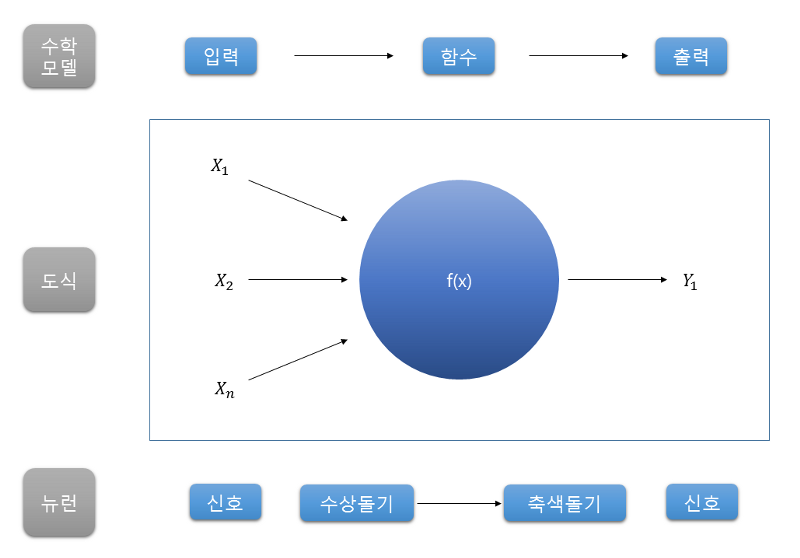
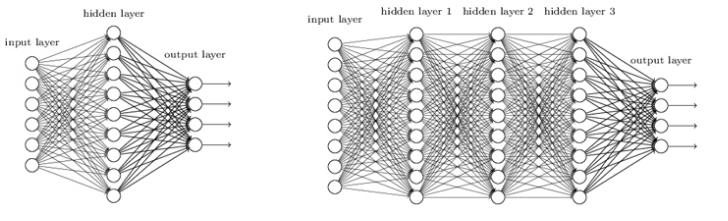
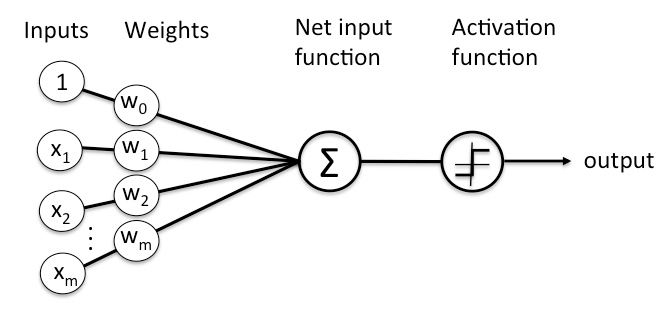
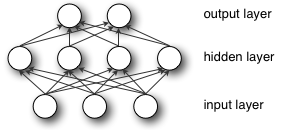
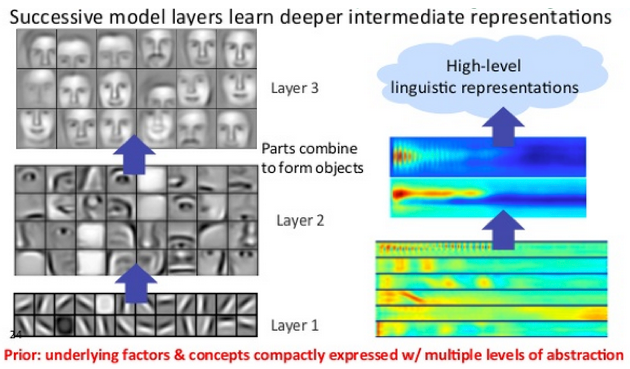
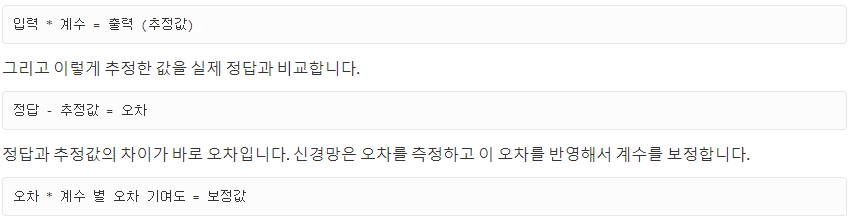
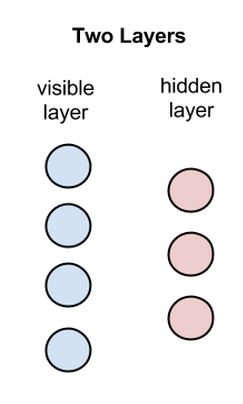
Ⅰ. 인공신경망

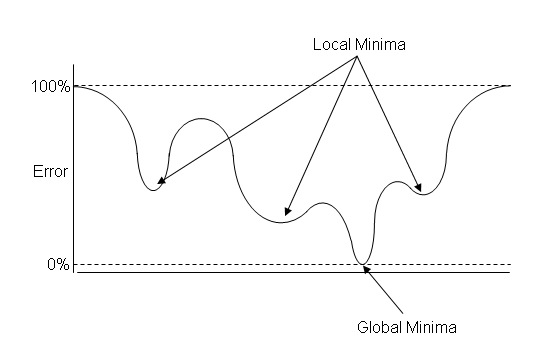
1. 왜? 
2. 
3. 
4. 그렇다면 이렇게 설계된 인공지능이 과연 인간의 지능과 유사할까?
   1. 인간은 기억을 잊을 수 있다
   2. 인간은 사건이 발생한 매 순간순간 배우지 않는다.
   3. 인간은 협소한 지식 또는 편견을 배울 수 있다.
   4. 인간은 감정을 갖는다.
   5. 인공지능은 모든 데이터를 외워버릴 수 있다 (과적합 문제 > Dropout or Batch Normalization 사용)
   6. 인공지능은 우물 안 개구리가 되어 빠져 나오지 못할 수 있다 (Local minima 문제 > backpropagation or pre-training 사용)
5. 인공신경망 구성 요소
   1. 심층 신경망은 여러 개의 층으로 이뤄진 신경망을 의미
   2. 한 층은 여러 개의 노드로 구성됨.
   3. 노드는 실제 연산이 일어나며, 일정 크기 이상의 자극을 받을 시 반응함. 그 반응의 크기는 입력값과 노드의 계수(or 가중치, Weight)를 곱한 값과 비례
   4. 이 계수를 조절하여 여러 입력에 다른 가중치를 부여할 수 있음.
   5. 최종적으로 곱한 값들을 전부 더하고 그 합을 활성함수의 입력에 들어감.
   6. 활성 함수의 결과가 최종 값이며, 이 출력 값이 분류나 회귀분석에 사용됨
   7. 
6. 머신러닝과 딥러닝의 차이점 : shallow vs. deep, 특징추출 자동 유무
7. 
8. 인공 신경망 학습 목적 : 출력의 오류를 최소화 하는 것
9. 인공 신경망 학습 과정 ;
   1. 학습이 시작되기 전 뉴럴넷의 모든 계수 초기화
   2. 반복적으로 데이터를 보여주어 학습
   3. 만일 학습이 잘 되었다면 계수는 적절한 값으로 업데이트. 이 신경망으로 각종 분류, 예측
   4. 학습 과정 내부에선 같은 원리의 계수 업데이트 반복 발생
   5. 계수 업데이트의 원리는 우선 계수를 추정하고 그 계수를 사용했을 때 발생하는 에러를 측정한 뒤 그 에러에 기반해 계수를 약간씩 업데이트 하는 것
   6. 수식
   7. 선형 회귀 (X는 입력, b는 기울기, a는 y절편, Y\_hat은 추정값)
   8. 다중 회귀 분석 (층의 출력 = 다음 단계 층의 입력, X\_1, X\_2, X\_3는 이 입력값, b\_1, b\_2, b\_3는 노드가 갖고 있는 계수로 입력 값을 얼마나 반영할 것인지 정하는 가중치)
      1. Y\_hat은 출력으로 나가기 전에 비선형 함수를 통과함. 그 사용 이유는 출력의 범위를 제한하여 노드의 출력값이 발산하지 않게 해줌. 제일 중요한 목적은 함수의 비선형성 부가)
      2. 최종 출력단을 0과 1로 제한할 경우, 특정 2가지에 대한 분류
      3. 로지스틱 회귀 분석을 출력단에 붙일 경우, 출력값은 0~1 연속적인 값으로
      4. 각 노드에 적용되는 비선형 함수는 S자 모양의 비선형성을 가짐. Sigmoid 함수나 하이퍼탄젠트 함수, 렐루 함수 등. 이 비선형 함수의 출력이 다음 층의 입력값이 됨.
   9. 계수 업데이트 방법 – Gradient Descent(경사 하강법)
      1. Gradient 는 경사나 기울기를 의미. X-y 평면에서 두 변수의 관계를 설명
      2. 계수를 어떻게 결정해야 전체 에러가 가장 작아질까? 어떻게 노드의 계수를 설정해야 입력 데이터에서 의미 있는 정보를 추출할 수 있을까? 어떻게 해야 사진의 물체가 사람인지, 자동차인지, 고양이인지 알아낼 수 있을까?
      3. 인공신경망의 학습과정은 수 많은 계수를 보정해가는 과정. 이 계수는 오차를 줄이는 방향으로 보정되어야 데이터에서 의미 있는 정보 추출 가능. 오차와 계수의 관계는 편미분으로 정의.
      4. 심층 신경망은 여러 층으로 이루어져 있기 때문에 각 노드의 계수는 신호에 곱해지는 여러 계수 중 하나. 그러므로 연쇄법칙을 사용해 설명
      5. 연쇄 법칙 , (오차와 계수가 활성 함수를 통해 정해지는 경우 최종적으로 계수 값의 변화가 오차를 증가할지, 혹은 감소할지 구할 수 있음. 활성함수를 바꾸는 것이 어떤 영향을 끼칠지 알 수 있음)
   10. 로지스틱 회귀 함수
       1. 심층 신경망의 여러 층 중 마지막 층, 즉 출력층은 다른 층과 조금 다름.
       2. 분류 작업의 경우 최종적으로 출력 노드는 0 또는 1의 값을 가짐.
       3. 입력 데이터는 출력과 달리 연속적인 값을 가짐
       4. 넓은 범위의 값을 0 또는 1의 값으로 결정해주는 과정을 로지스틱 회귀라 함. 이 회귀는 이름은 회귀지만 실제로 분류 작업을 함
       5. 수식
          1. 연속된 값(-100, -0.5, 1.0, e^4, 300, …)을 확률로 표현하려면 입력값을 양수로 바꿔야 함(확률은 0~1 사이) 이런 조작을 위해 분모에 e^(-x) 가 들어감. X에 어떤 값이 들어오더라도 결과는 항상 양수
          2. 입력값 x가 클 경우 확률이 점점 증가하도록 설계해야 함. 그래서 e^(-x)가 분모에 포함되어 있음. 만일 x가 점점 커지면 e^(-x)는 점점 작아짐. F(x)는 1에 가까워짐
          3. 지수 함수로 처리해야 결과를 0과 1사이로 압축시키기 쉬움.
          4. 확률의 기준값을 세워주어, 그 값을 통해 분류 작업 수행

RBM

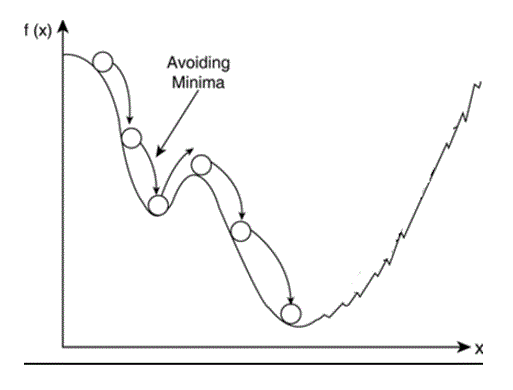
1. 정의와 구조 : 제한된 볼츠만 머신
   1. 차원 감소, 분류, 선형 회귀 분석, 협업 필터링, 특징값 학습, 주제 모델링
   2. RBM 자체로 쓰이기도 하지만 주로 다른 심층 신경망의 학습을 도움
   3. 구조 : 2개의 층(입력층 1개, 은닉층 1개) > 심층 신경망이 아님
      1. 가시층 : 데이터가 입력되는 층
      2. 은닉층 : 특징값이 학습되는 층
      3. 모든 은닉층의 노드는 모든 가시층과 연결되어 있음
      4. 같은 레이어에 속한 노드는 서로 연결X



* + 1. 심층 신뢰 신경망(DBN:Deep Belief Network)을 구성하는 요소로 쓰임
    2. 같은 층 내부 연결이 전혀 없으며 RBM이라 불리는 이유.
    3. 입력층의 노드는 데이터를 입력받으며 입력받은 데이터를 은닉층에 얼마나 전달할지 확률에 따라 결정(stochastic decision, 확률적 경사 하강법) 확률에 따라 전달할지(1) or 전달하지 않을지(0)
    4. 확률적 경사 하강법 : 딥러닝의 속도 개선을 위한 방법. 지역 최소값에 빠질 위험 발생



* + 1. 모멘텀 : 지역 최소값에 빠질 위험을 없애기 위해 개선한 방법.



공에 관성을 부여하여 작은 언덕을 넘게 하는 기술

기존에 업데이트에 사용했던 기울기의 일정 %를 남겨 현재의 기울기와 더하여 업데이트에 사용하는 방식