

\*\*\*\*\*딥러닝 응용\*\*\*\*\*

#1주차

공지

연강하면 12시 20분(12시 10분까지 수업, 나머지 10분은 질문타임)

중간고사는 인공지능 기초를 위한 FAQ 내용이 포함된다.

수업내용

pre-trained: 사전 훈련

반지도학습(semi-supervise)->비지도학습의 반, 지도학습의 반을 섞은거

전통적인 :input,output을 주고 실행

ai: 룰을 주고 실행

머신러닝:(입력과 룰을 주고 출력을 얻어낸다.)=(데이터로부터 결과를 얻는다.)

확률론으로 볼때:

데이터가 정규분포를 따를때 -> 안따를때보다 더 분류를 할수 있다.(머신러닝)

" " 안따를때 -> 분포를 만들어 내는 함수를 구할수 있다.(딥러닝)

머신러닝과 딥러닝의 차이점: 픽처(특성)를 추출하냐 안하냐의 차이점이다.

입력->모델->훈련->출력

이때 훈련 할때 픽처를 추출한다.(딥러닝)

머신러닝은 입력하기전에 추출한다.(머신러닝)

classification, regression 의 차이점: 분류(둘 이상의 특징을 서로 나눔) ,회귀(연속적인 값을 정리)

classification을 잘게 나누면 결국 regression이 된다.

feature가 많을 수록 분류가 잘됨.

차원이 많을수록 weight하고 bias를 뽑기가 힘들다.->과적합(overfitting)

해결방법:픽처의 중요도(feature selection) 를 계산하다. 규제(regularization)를 한다.

weight를 0으로 근사하게 설정한다.

- 학습의 뜻:weight하고 bias를 얻어서 계산해 결과를 추출함.

overfitting과 underfitting의 차이점: overfitting은 많은 데이터(noise와 잘못된 데이터)이 있어 잘못된 결과를 추출 하는 경우 underfitting: 데이터를 너무 적게 뽑아서 생긴 문제

chat gpt의 파라미터의 갯수:1750억

딥러닝

1.활성화 함수:층을 쌓으면 쌓을수록 특징이 안쌓인다. 비선형 함수를 사용해서 한번 꼬아준다.

2.옵티마이저: 경사하강법(최적의 값을 찾아가는 방법)

3.백프로파게이션:

4.포워드프로파게이션:

5.로스트 평선: 측정값과 예상값의 차이점을 계산하는 함수(cce,bce,mse)

6.원핫 인코딩: 0,1로 변형해서함. 숫자 사이의 관계 땀에 사용됨

경사하강법

$$W_{t+1}=W_t-rG$$

r->running mate(학습률)

G->현재 미분값,기울기 값

층이 많이 쌓을수록 백프로파게이션 할때 기울기가 소실된다.(vanishing gradient,기울기 소실문제)

->rnn,lstm,transformers가 나옴.