

2025-1 인공지능과 뇌인지과학 기말고사 정리 (20241519 조예성)

머신러닝 시스템 분류 기준 3가지: 훈련지도여부 / 실시간 훈련 여부 / 예측 모델 사용 여부

- 예측 모델 사용 유무: 사례 기반 학습, 모델 기반 학습
- 훈련 지도 유무: 지도학습, 비지도학습, 강화학습
- 실시간 유무: 실시간 학습, 배치 학습

인공지능의 비 베타성: 여러 학습 기반을 동시에 가질 수 있다

- 모델기반학습, 온라인 학습, 지도학습

답을 붙이는 작업: 레이블링

예측 변수라 부르는 특성을 사용해 타깃 수치를 예측하는 작업: 회귀

MSE를 구하시오:

실제 값	예측 값	오차	MSE
3	2		
7	10		

비지도 학습 종류: 계층 군집, 시각화, 차원 축소, 이상치 탐지, 특이치 탐지, 연관 규칙 학습

시각화 다른 말: EDA

차원 축소 내부: 특성 추출 feature extraction

에이전트에게 보상 또는 벌점을 주어 가장 큰 보상을 유도하는 학습: 강화학습

주어진 세트 전체를 활용해 오프라인에서 훈련: 배치학습

- 단점: 컴퓨팅 자원 많이 필요, 새로운 데이터 입력시 처음부터 학습

미니배치를 사용해 점진적으로 훈련: 온라인 학습

- 단점: 나쁜 데이터가 주입되는 경우, 시스템 성능 하락, 지속적 모니터 필요
- 데이터 적응 속도: 학습률

사례기반 학습 대표 예시: KNN 알고리즘

- K는 항상 (홀수)다

모델 기반 학습 대표: 선형 회귀

알고리즘 학습 모델에 사용되고, 훈련 과정 중 변하는 값이 아닌 값은? 하이퍼파라미터

적절한 모델을 알기 위하여서는 직접 튜닝을 통해 확인해야 하는 점을 칭하는 용어는? 공자 점심 없음

인공지능 에이전트를 여러 개 엮어 사용: Agent to Agent

행렬 계산을 강력한 라이브러리: numpy

MSE 식

오차를 다른 용어로: 비용

모델이 얼마나 나쁜지 계산하는 함수: 비용함수

비용 함수가 가질 수 있는 최솟값: 전역 최솟값

훈련 과정에서의 비용함수 파라미터 조정 비율: 학습률

경사를 타고 내려오는 방법: 경사 하강법

방향 결정: 미분한 값

학습률이 너무 큰 경우: 비용함수 수렴하지 않을 가능성 있음

비용이 가장 낮은 세타를 알아내기: 경사하강법

항상 최솟값에 도달한다: O, X

최솟값에 도달하는 조건: 2차 함수

이상한 기울기 0: 지역 최솟값

학습률: 학습스텝

특성을 0~1 사이 값으로 변환: 특성 스케일링

마진 폭을 최대로 하는 분류: SVM Support Vector Machine

도로 안 샘플 허용 여부: 하드마진 분류, 소프트 마진 분류

SVM 단점: 훈련 세트가 선형적으로 구분되는 경우에만 가능

모든 훈련 샘플이 도로 바깥쪽에 올바르게 분류되도록 하는 분류: 하드 마진 분류

C는 클수록 마진 오류가 (적게), 도로폭을 (작게) 만듦, c를 info로 지정하면 (하드 마진 분류)

하이퍼파라미터의 조합을 평가하여 최적의 선택 결정: Grid Search

예측값과 실제 라벨 사이의 차이카 클수록 큰 손실이 가해지는 것: hinge 손실

쌍대성 허용여부 : dual

뇌에 있는 생물학적 뉴런의 네트워크에서 영감을 받은 머신러닝 모델: ANN

최초의 인공신경망 만든 2명: 워런 매컬러, 월터 피트

- 각자 전공: 신경생리학자, 수학자

인공지능 1차 침체기 1960년

1990년대 신규 개발 된 기술: SVM

인공 신경망 부흥 이유 3개: 데이터 많아짐, 하드웨어 발전, 알고리즘 향상

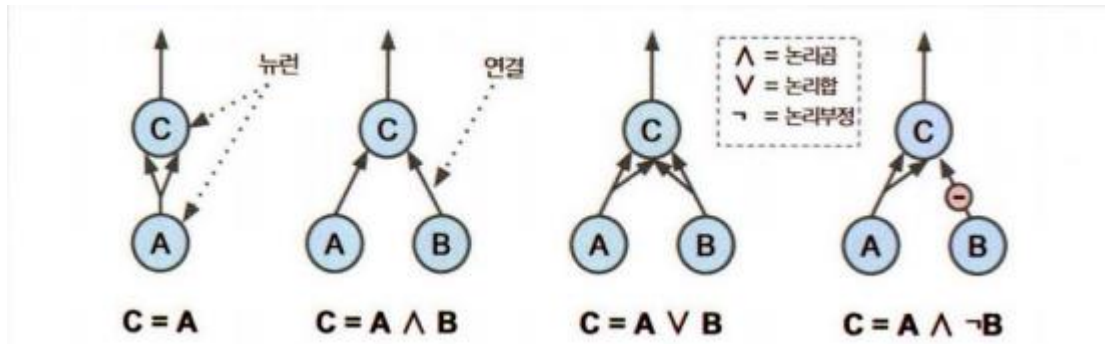
이론에서보다 큰 문제가 안된 문제: 지역 최저점 문제

뉴런의 구조 7개: 수상돌기, 축삭, 축삭 종말, 시냅스, 활동 전위, 신호, 신경전달 물질

인공 뉴런은 하나 이상의 (이진) 입력과 (이진) 출력을 가짐

출력 내보내는 조건: 입력이 일정 개수만큼 활성화 되었을 때

퍼셉트론 다른 이름: TLU(Threshold Logic Unit); 임계값 로직 유닛



계단 함수 적용 위치: 입력값과 가중치를 곱한 값들의 합에 적용

가장 많이 사용되는 계단 함수: Heaviside, sign 함수

하나의 TLU를 간단한 이진 분류기로 활용이 가능하다 (O , X)

편향값 1을 항상 출력하는 편향 뉴런과 입력 뉴런이 함께 사용되는 층: 입력층

퍼셉트론을 여러 개 쌓아올린 인공신경망: 다층 퍼셉트론

층에 속한 각각의 뉴런이 이전 층의 모든 뉴런과 연결되어 있을 때를 가리키는 말: 완전연결 층

하나의 층에서 이루어지는 입;출력을 행렬 수식으로: $h_{w,b}(X) = \phi(XW + b)$

여러 개의 은닉층을 쌓아올린 인공신경망 입력 샘플들의 특성 행렬: 심층신경망

Barkpropagation 알고리즘 탄생: 1986년

GPT 풀 언어: Generative Pre-Trained Transformer

자연어 처리 영어: NLP, Natural Language Processing

AI업계는 (좁은 AI) 에서 (다목적 범용 AI)로 dlehdgkduTek

텍스트 생성 AI: GPT, BERT, LLaMA

이미지 생성 AI: DALL-E, StyleGAN

음성 생성: WaveNet, Tacotron

영상 생성: Sora

2010년대 주요 딥러닝 모델 2가지: CNN, RNN

컴퓨터 비전: Cnn

자연어 처리: RNN

각 단어들 간의 중요도를 계산: Attention

RNN은 (병렬) 처리가 불가하다

언어 이해 및 생성 능력을 극대화한 모델: LLM

- 하이퍼 파라미터 6개: epochs, Learning Rate, Batch Size, Max Output Tokens, Decoding Type, Top-k & Top-p 샘플링

텍스트를 작은 단위로 나눈 것: 토큰

단어나 문장을 고정된 크기의 벡터로 변환하는 과정: Embedding

- 어떠한 저주 해결: 차원의 저주

LLM에서 디코딩이란? 응답 생성

매 단계에서 가장 높은 확률을 가진 토큰 선택: 그리디 디코딩

여러 경로를 동시에 탐색하며 가장 높은 누적 확률 가진 경로 선택: 빔 디코딩

각 단계에서 확률 분포에 따라 토큰을 무작위로 선택하여 응답 생성: 샘플링 디코딩

확률순 k개 선택하여 그중 출력: Top-k

누적 확률이 p에 도달할(이하) 때까지 토큰 선택: Top-p

둘중 더 많이 사용하는 것: Top-P

LLM 활용 앱 개발을 위한 오픈소스 프레임워크: 랭체인 LangChain

LLM이 기존에 학습하지 않은 새로운 데이터를 유연하게 활용할 수 있게 해주는 시스템: VectorDB

- 어떤 것을 기준으로 검색: 유사도
- 텍스트, 이미지, 오디오에 사용 가능: O