|  |  |
| --- | --- |
| 교육 제목 | AI 프로젝트 기반 S/W 전문가 양성과정 |
| 교육 일시 | 21.11.22 |
| 교육 장소 | 대면(강의실) |
| **교육 내용** | |
| 오전 | 회귀를 위한 기초 수학  일차 함수, 기울기와 y절편  이차 함수와 최솟값  \*미분, 순간 변화율과 기울기  편미분  지수와 지수 함수  시그모이드 함수  로그와 로그 함수  가장 훌륭한 예측선 긋기: 선형 회귀  선형 회귀의 정의  가장 훌륭한 예측선이란?  최소 제곱법  평균 제곱 오차  잘못 그은 선 바로잡기  오차 수정하기: 경사하강 법  경사 하강법의 개요  학습률  다중 선형 회귀란  참 거짓 판단 장치: 로지스틱 회귀  로지스틱 회귀의 정의  시그모이드 함수  오차 공식  로그 함수  로지스틱 회귀에서 퍼셉트론으로  새로운 데이터 받음 -> 전처리 방식 선택위한 작업  데이터 타입 파악,  데이터 탐색,  그래프 그리기 |
| 오후 | 혼자 공부하는 머신러닝 딥러닝  마켓과 머신러닝  k-최근접 이웃 알고리즘 : 가장 간단한 머신러닝 알고리즘 중 하나.  전체 데이터를 메모리에 가지고 있는 것이 전부  머신러닝에서의 모델 : 알고리즘이 구현이 된 상태  훈련 : 머신러닝 알고리즘이 데이터에서 규칙을 찾는 과정  정확도 : 정확한답을 맞춘 비율, 정확도 = hit count / all data count  훈련 세트와 테스트 세트  지도학습과 비지도학습  지도학습 : 훈련하기 위해 데이터와 정답이 필요  비지도학습 : 스스로 데이터를 찾아서 학습  훈련세트와 평가세트 : 머신러닝의 모델을 만들기 위해 사용 하는게  훈련세트, 모델의 정확도를 테스트 하기 위해 사용 하는게 평가 세트  셈플링 편향 : 훈련 세트와 평가 세트샘플이 골고루 섞여 있지 않으면  데이터 셈플링이 한쪽으로 치우칠 상황이 생김  훈련모델 평가 : 모델을 훈련할 때 사용한 데이터로 평가하면 정답을 미리  알려주는 꼴이므로 훈련에 참여하지 않은 데이터로 평가 해야 한다  데이터 전처리  Numpy로 데이터 준비  Column\_stack()  Concatenate()  사이킷런으로 훈련세트와 테스트세트 나누기  Train\_test\_split()  기준 맞추기  데이터 전처리 : 알고리즘들의 셈플들 간 거리에 영향을 덜 받게 하기  위해 특성값을 일정한 기준으로 정렬헤야 한다.  데이터 전처리 방법 : 표준점수  브로드캐스팅  크기가 다른 넘파이 배열에서 자동으로 사칙 연산을 모든 행이나  열로 확장하여 수행하는 기능  K-최근접 이웃 회귀  K-최근접 이웃 회귀  지도학습 알고리즘 : 분류, 회귀  회귀 : 클래스 중 하나로 분류하는 것이 아닌 임의의 수치를 출력  결정계수 (R^2)  사이킷런에서 K-최근접 이숫 회귀 알고리즘을 구현한 클래스  :KNeighborsRegressor  과대적합 vs과소적합  과대적합 : 훈련에서는 좋았는데 테스트세트에서 점수가 굉장히 낮음  과소적합 : 훈련세트보다 테스트데이터 성적이 더 좋음  회귀문제 다루기  사이킷런은 회귀모델 점수로 R^2(결정계수)값을 반환, 이값이 1에  가까울수록 좋음, 이때 모델을 훈련하고 나면 과소적합이나 과대적합이  나올 수가 있다  선형회귀  K-최근접 이웃의 한계  시간과 환경이 변화하기 때문에 주기적으로 새모델로 훈련 시켜야 한다.  선형회귀  널리 사용중인 회귀 알고리즘 중 하나. 비교적 간단하고 성능이 뛰어남.  다항회귀  선형회귀에 비해 다항식을 사용하기 때문에 좀더 훈련세트의 경향을 따라감  모델 파라미터  선형회귀가 찾은 가중치처럼 머신러닝 모델이 특성에서 학습한 파라미터  특성 공학과 규제  다중회귀 여러 개의 특성을 사용한 선형회귀를 다중회귀  특성공학  기존의 특성을 사용해 새로운 특성을 뽑아내는 작업  사이킷런의 변환기  사이킷런은 특성을 만들거나 전처리하기 위한 다양한 클래스를 제공  Fit(), score(), predict() 메서드  규제  머신러닝 모델이 훈련세트를 너무 과도하게 학습하지 못하도록  훼방하는 것을 지칭, 선형모델인 경우 특성에 곱해지는 계수의 크기를  작게 만드는 일  릿지회귀  규제가 있는 선형모델 중 하나, 선형 모델의 계수를 작게 만들어 과대적합  완화.  라쏘  또 다른 규제가 있는 선형회귀모델 |