|  |  |
| --- | --- |
| 교육 제목 | AI 프로젝트 기반 S/W 전문가 양성과정 |
| 교육 일시 | 21.12.2 |
| 교육 장소 | YGL |
| **교육 내용** | |
| 오전 | 1. 모델 REVIEW 데이터 가져와서 정렬하고 어떤 모델 쓸지 설정하고 층을 얹을지 설정하고 컴파일 = 손실함수 fit 2. 데이터 다루기 pima Indian data – 당뇨가 유전인지 아닌지 질이 좋은 데이터 set이 있어야 함 – 분석 성공의 요인 전처리(정제) 🡪 트레인, 테스트(검증) 나누기 🡪 예측 3. 실습 – pima Indian 데이터 describe로 상관관계 확인 4. 실습 – 딥러닝 5. 다중분류 – softmax, 손실함수 = crossentropy 다중분류는 pairplot을 자주 씀 6. 문자열을 숫자로 변환 from sklearn.preprocessing import LabelEncoder e = LabelEncoder() e.fit(Y\_obj) Y = e.transform(Y\_obj) 7. to\_categorical 사용 공부하기 8. 실습데이터 초음파 광물 예측 – 과적합 학습 set, 테스트 set 나누기 from sklearn.model\_selection import train\_test\_split X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size= 0.3, random\_state = 0) 9. 입력 요소가 많으면 노드 수를 높임 차원이 높으면 노드를 줄이고 레이어를 늘리고 10. input\_dim=X.shape[1] |
| 오후 | 1. 딥러닝 – 미니 프로젝트 고민 해보기 2. train, validation, test 3. 재사용하기 위해 저장, 사용 model.save(“파일명.h5”) 4. 저장한 모델 불러오기 from keras.models import load\_model model = load\_model(\*.h5) 5. 교차 검증 cross validation    1. K-fold cross validation (K겹 교차검증) k개로 쪼개고 쪼갠 것을 train, test로 나누고 각각 계산 결과를 평균 냄 6. batch\_size가 줄어들면 속도는 줄어들고 정확도는 높아짐 7. 참고 class A: \_\_call\_\_(self):  print(call) 8. 베스트 모델 만들기 - loss가 줄지 않으면 멈추게 하기 df.sample(frac=1) 전체를 전부 sample로 가져오기 from keras.callbacks import ModelCheckpoint, EarlyStopping checkpointer = ModelCheckpoint(modelpath, monitor = 'val\_loss',  verbose = 1, save\_best\_only = True)   verbose = 게이지, save\_best\_only = True 좋아지면 save 9. 정리    1. 데이터 셋 확인    2. 데이터 정제 - 결측치 처리 - 문자를 숫자로 변환 – LabelEncoder 🡪 transform - 원-핫 인코딩 – keras.to\_categorical(Y)    3. 모델 설정 - 어떤 모델을 사용할건지, Sequential() Dense - layer는 몇 개층, 어떤 층을 깔아줄지 - 출력은 몇 개 node 인지, 어떤 함수를 써줄지    4. 모델 compile 설정 - 손실함수(loss), optimizer, metrics…    5. 모델 fit (실행) - 몇 번 실행할 건지 epoch 가져오는 데이터 개수(batch\_size) - checkpoint 쓸껀지    6. 베스트 모델 만들기 - checkpoint, |