

<0318 인공지능개론>

• Data Loading

• Data Cleaning (EDA: Exploratory Data Analysis)

— 결측치 제거

— column이 필요 있는지 확인 후 삭제

— 대체 값으로 대체 가능한지 확인 (mean 등)

— Outlier 제거

— data imbalance 확인 후 outlier 제거

— Seaborn.boxplot() : 박스 그래프

— pyplot.scatter() : 산점도

• Data Scaling

— StandardScaler()

— MinmaxScaler()

• Diabetes data로 Data Loading, Data Cleaning, Data Scaling 추가해 몰리기

— 결측치 확인 및 제거

— Scatter

— Boxplot

— Labeling (갯수[value.count(), Graph] 추가

— Scaling (Standard, Min-Max) 비교