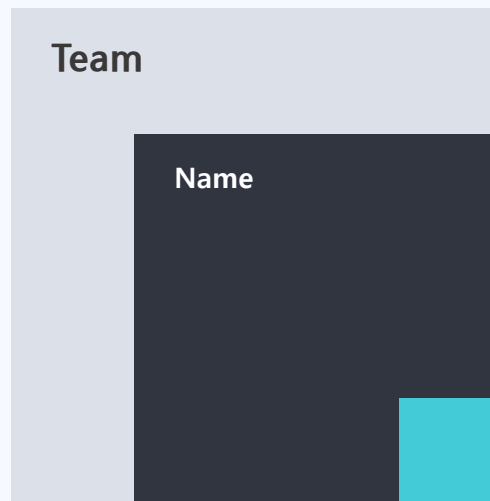


2018 DU BIGDATA CAMP

29. JUNE - 4. JULY

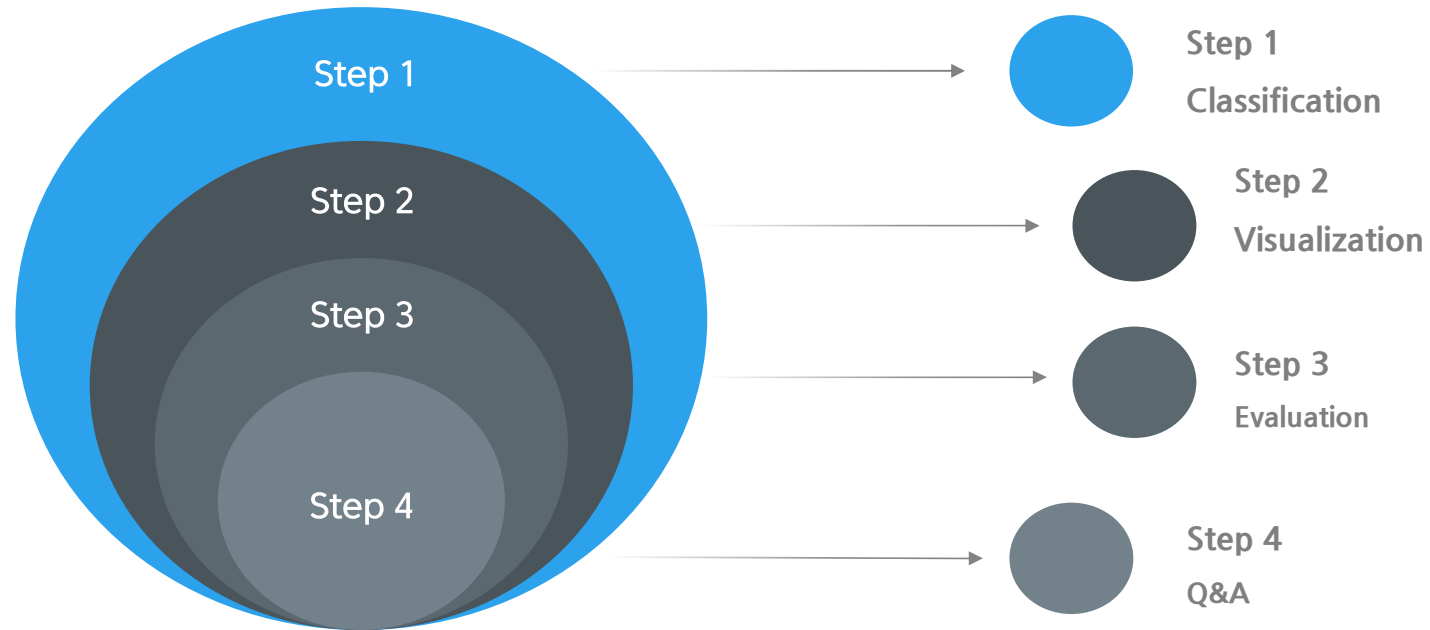


• 집에 언제가조

• 남승균 박상민
손주호 한해든
김소연 윤한나
김지범 김민제
안수빈

2018 DU BIGDATA CAMP

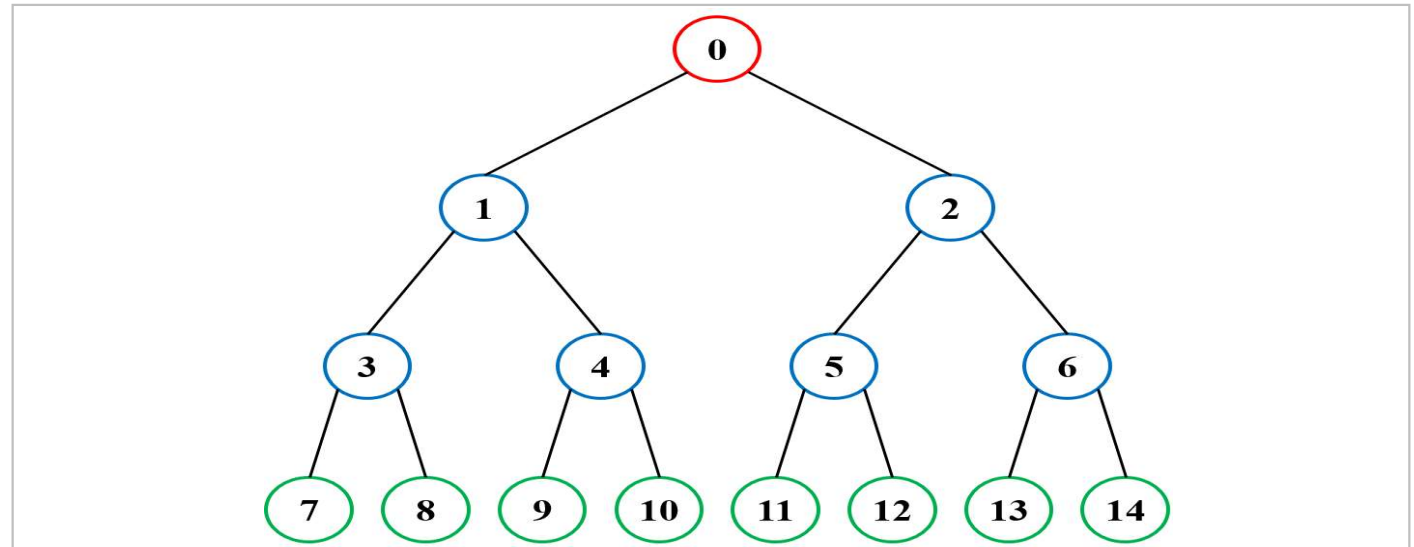
집에 언제가조



Classification

▶ Classification

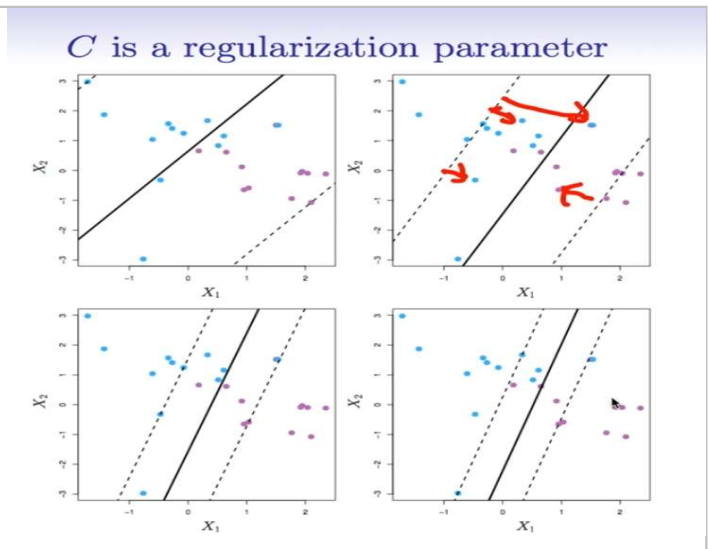
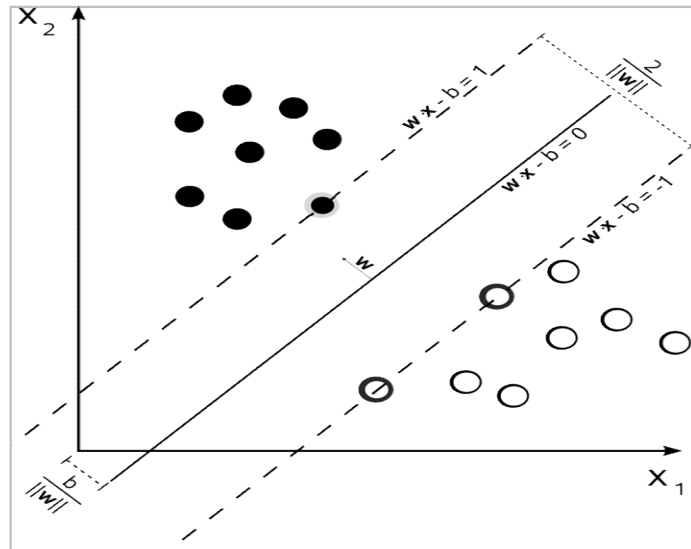
▶ Random Forest



노드, 결정 트리, 에지, 평균 예측치

▶ Classification

Support Vector Machine

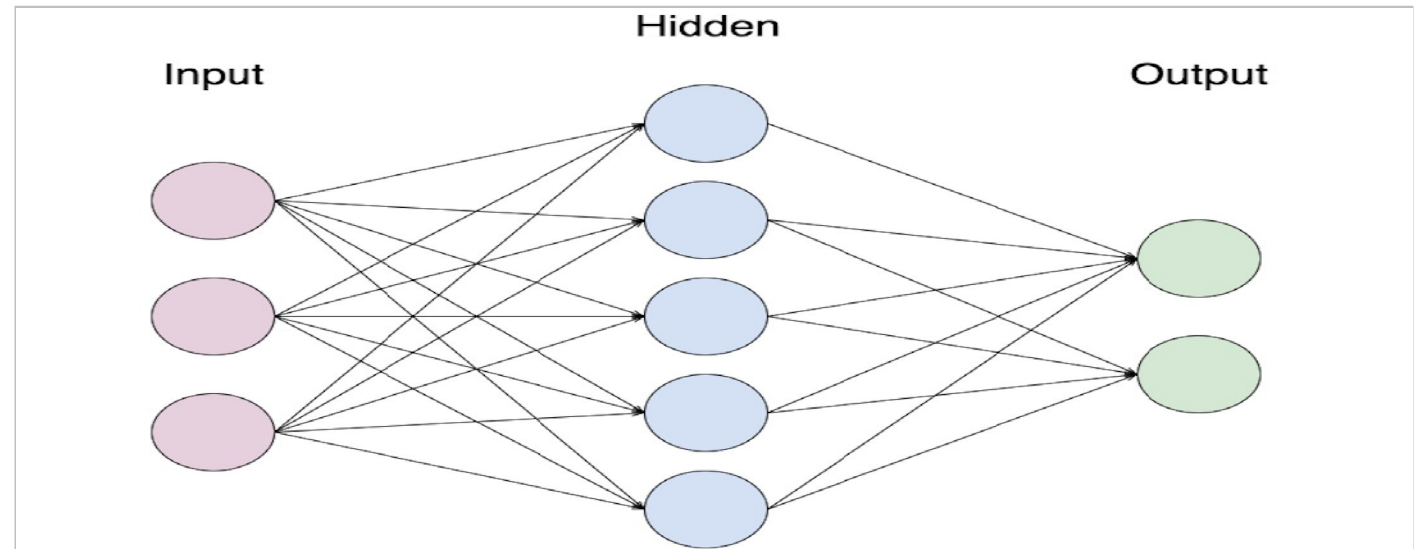


직선, 초평면, Margin



▶ Classification

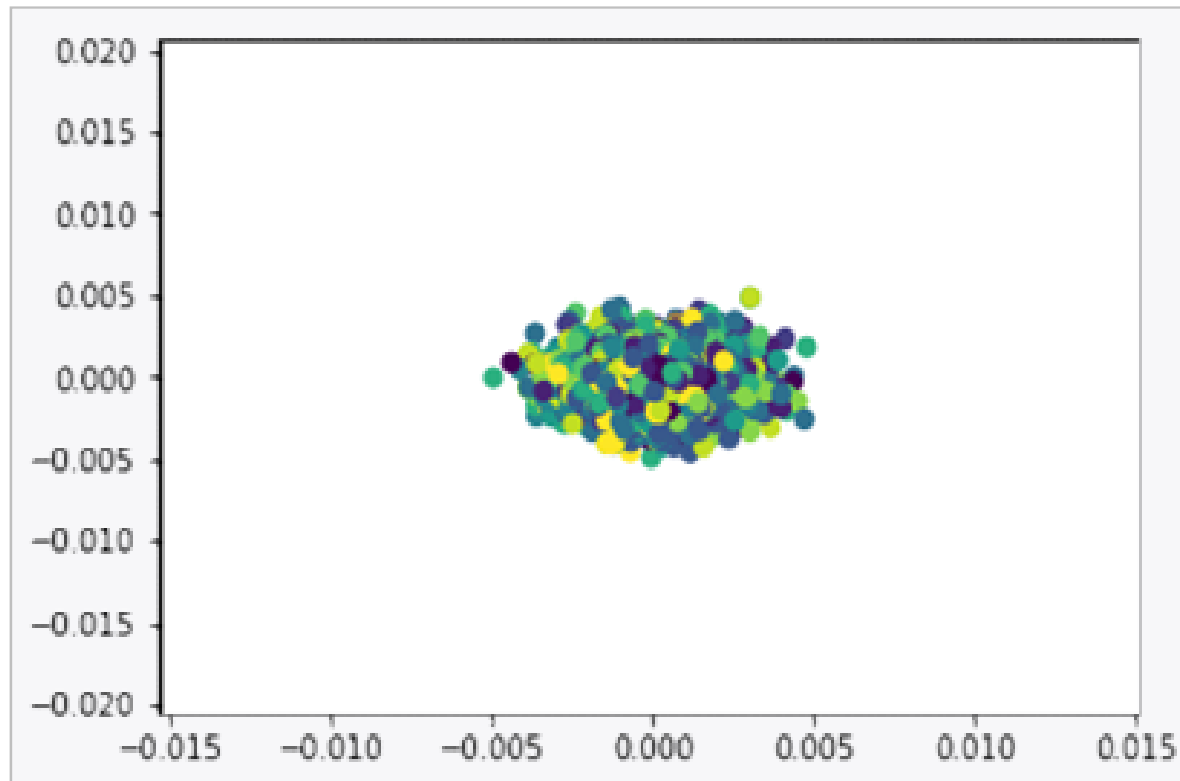
MultiLayer Perceptron



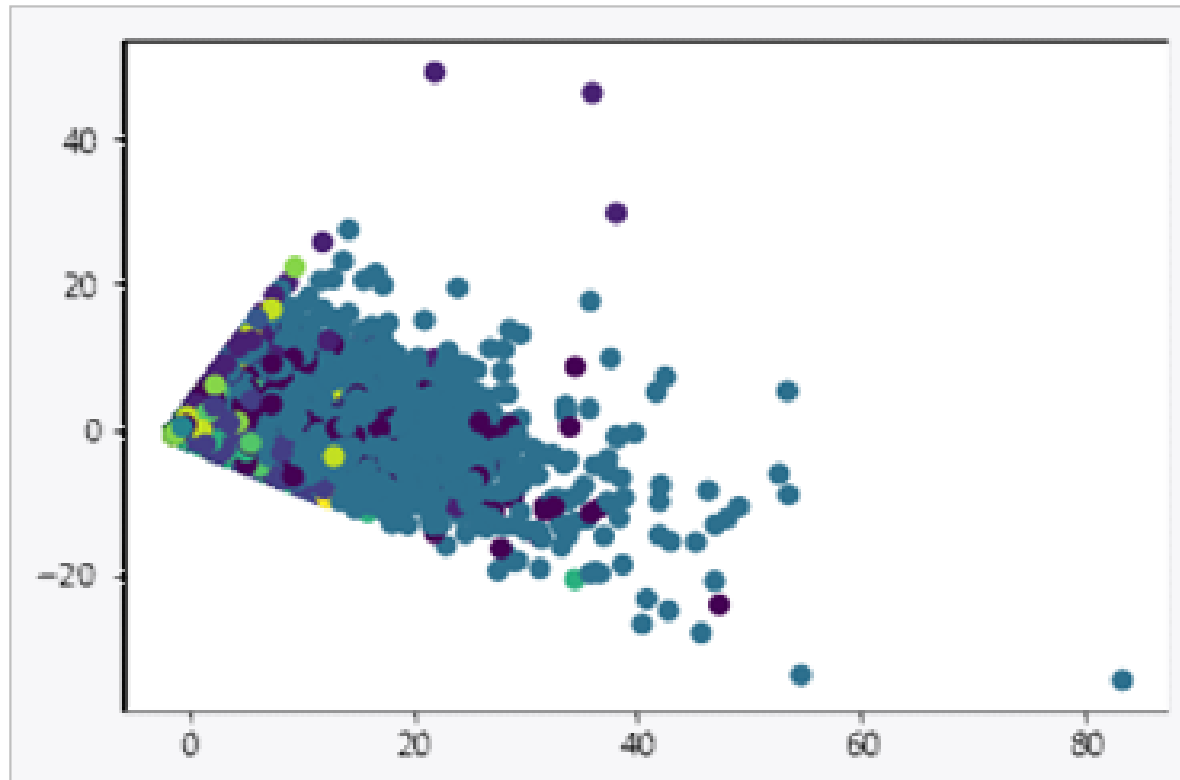
First(input) layer, second(hidden) layer, third(output) layer

Visualization

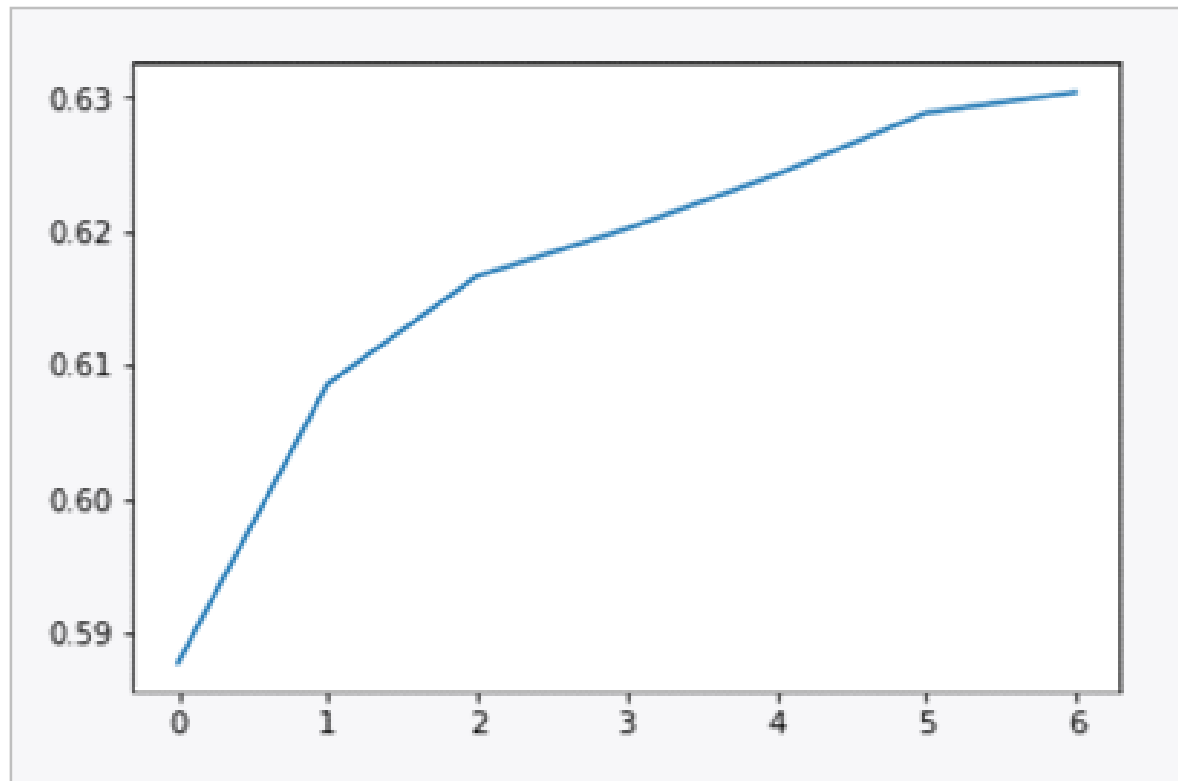
▶ t-SNE



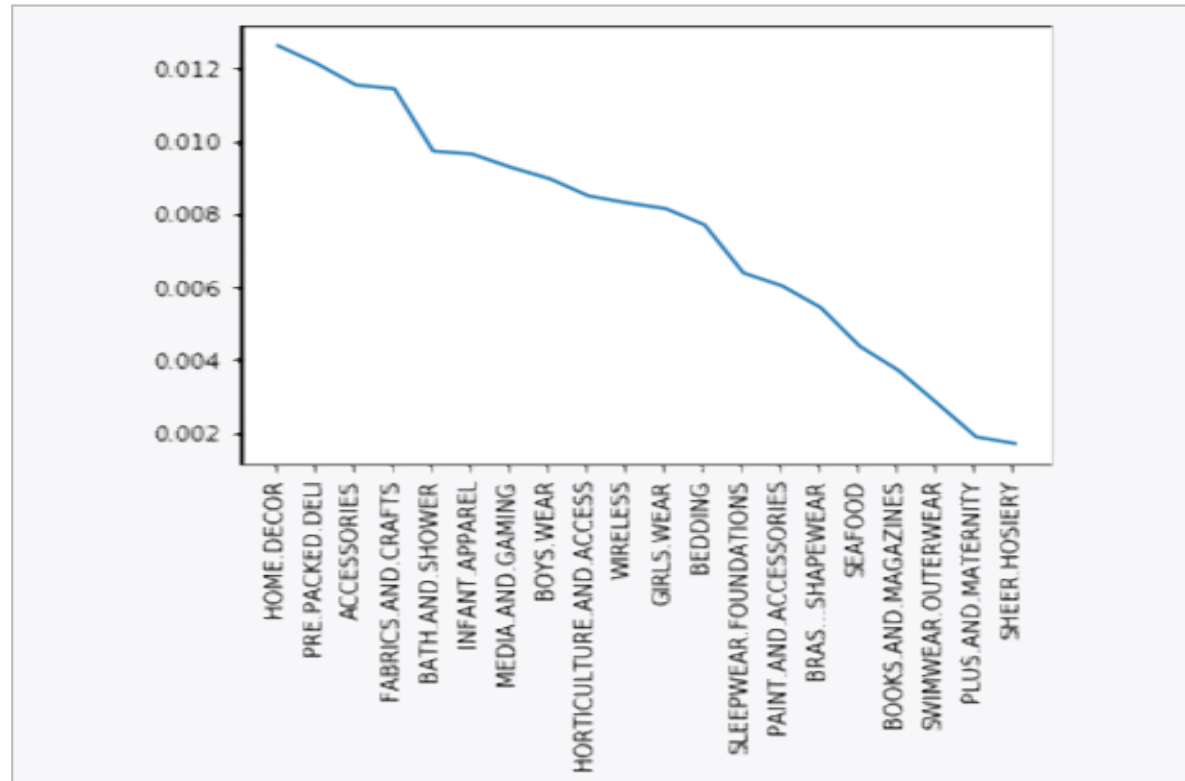
▶ PCA



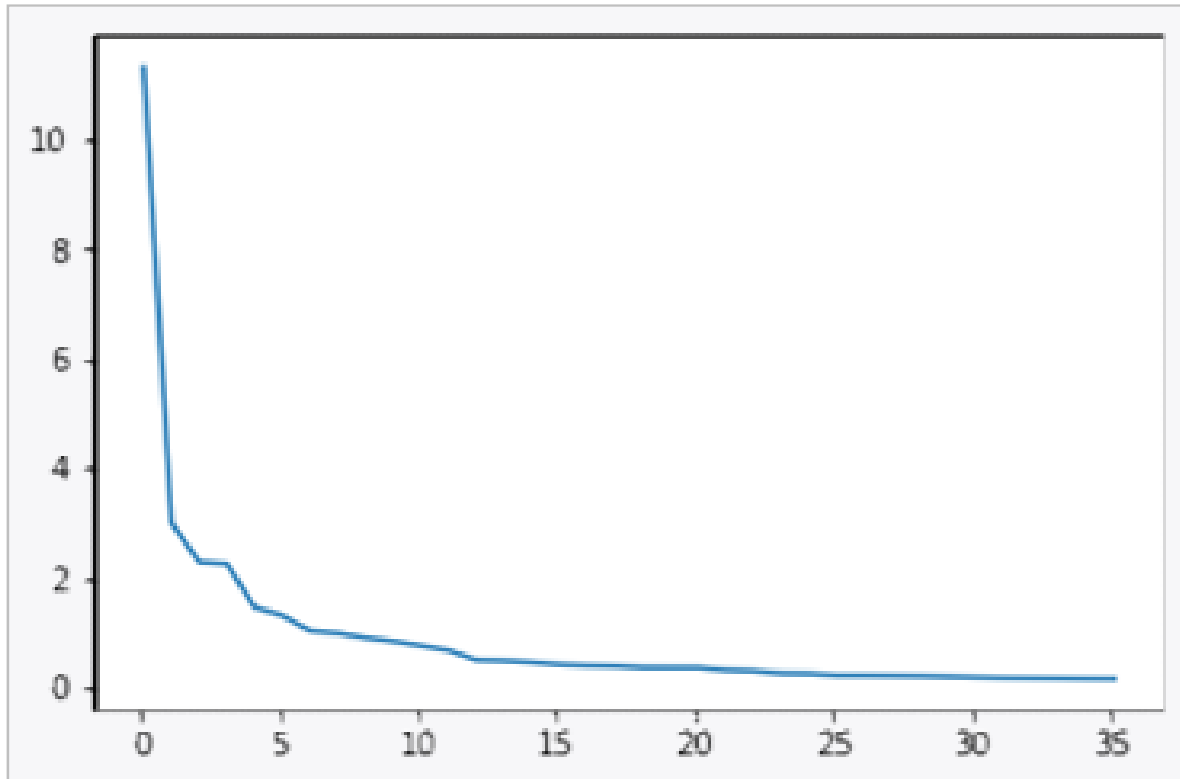
▶ Random Forest - TRAIN DATA



▶ RF 변수 중요도



▶ PCA 성분 수 결정



Evaluation

▶ Random Forest

n_estimators=2000
max_depth=200
max_features="log2"
min_samples_split=10

➔ 63%

▶ Support Vector Classification

C = 7
Kernel = 'rbf'

➔ 65.8%

▶ Neural Network

첫 번째 Layer : Dense=512
input dimension=40

두 번째 Layer : Dense = 256
input dimension=512

세 번째 Layer : Dense = 1
input dimension=256

activation = RELU
optimizer = adam
loss = mae
epochs = 5
batch_size = 64

➔ ? %

2018 DU BIGDATA CAMP



Q&A

