02과 딥러닝 개념 익히기

1.환경준비

(1) 라이브러리 로딩

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.metrics import *
from sklearn.preprocessing import MinMaxScaler

from keras.models import Sequential
from keras.layers import Dense
from keras.backend import clear_session
from keras.optimizers import Adam
from keras.callbacks import ModelCheckpoint
```

- 학습곡선 그래프

(2) 데이터로딩

```
path = 'https://raw.githubusercontent.com/DA4BAM/dataset/master/boston.csv'
data = pd.read_csv(path)
data.head()
```

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio	lstat	medv
0	0.00632	18.0	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3	4.98	24.0
1	0.02731	0.0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	9.14	21.6
2	0.02729	0.0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	4.03	34.7
3	0.03237	0.0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	2.94	33.4
4	0.06905	0.0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	5.33	36.2

Next steps: Generate code with data View recommended plots

변수	설명
medv	타운별 집값(중위수)
crim	범죄율
zn	25,000 평방피트를 초과 거주지역 비율
indus	비소매상업지역 면적 비율
chas	찰스강변 위치(범주 : 강변1, 아니면 0)
nox	일산화질소 농도
rm	주택당 방 수
age	1940년 이전에 건축된 주택의 비율
dis	직업센터의 거리
rad	방사형 고속도로까지의 거리
tax	재산세율
ptratio	학생/교사 비율
lstat	인구 중 하위 계층 비율

코딩을 시작하거나 AI로 코드를 생성하세요.

2.학습 절차 익히기

(1) 전처리

```
# 학습을 위한 데이터 분할
target = 'medv'
features = ['lstat']
x = data.loc[:, features]
y = data.loc[:, target]
x_train, x_val, y_train, y_val = train_test_split(x, y, test_size=.2, random_state = 20)

# 스케일링
scaler = MinMaxScaler()
x_train_s = scaler.fit_transform(x_train)
y_train_s = (y_train - y_train.min())/(y_train.max() - y_train.min())
```

(2) 학습절차 시각화 함수

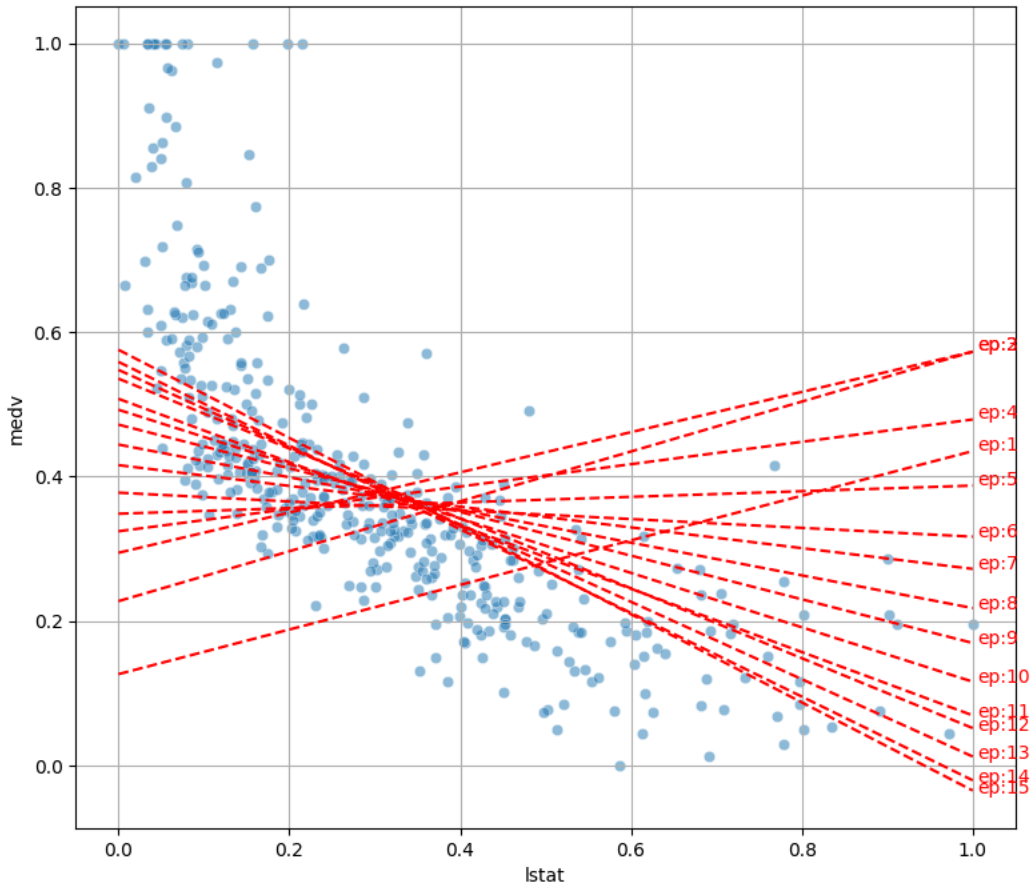
[ ] *숨겨진 셀 2개*

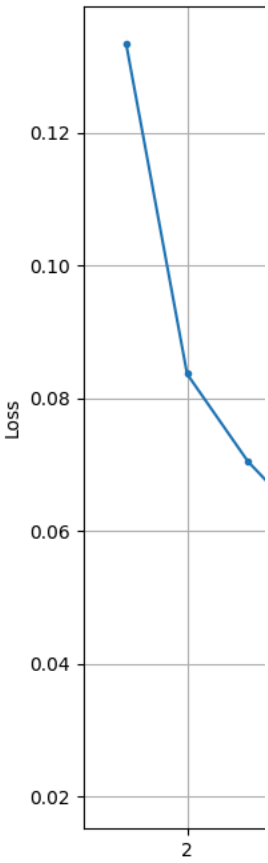
(3) 학습 절차 눈으로 확인하기

1) 실험1

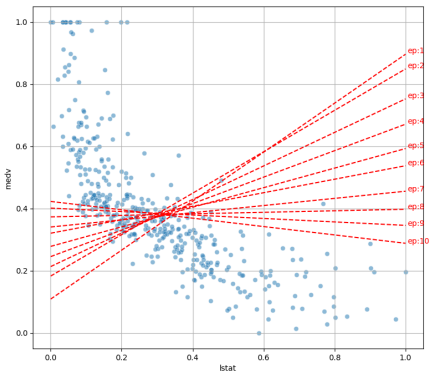
- epochs를 조절해 봅시다.
  - lr = 0.01로 고정시켜 놓고
  - ep를 5, 10, 15 로 조정하며 실행해 봅시다.

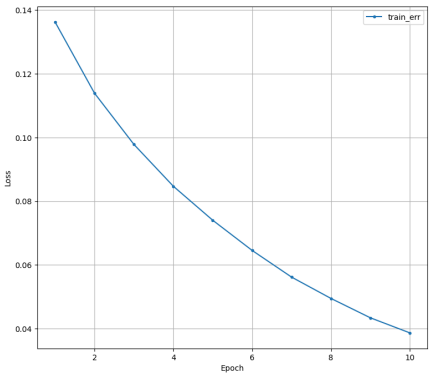
```
dl_visualize(ep = 15, lr = 0.01) # epochs : 학습 데이터를 몇번 학습 시킬거야? (15)
                                 # epochs 그래프 : 학습 곡선 엘보우보다 좀더 있는 곳을 선택
```



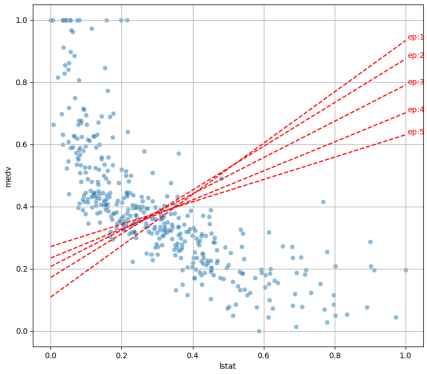


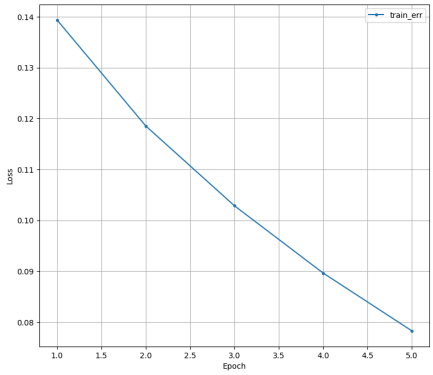

```
dl_visualize(ep = 10, lr = 0.01)
```






```
dl_visualize(ep = 5, lr = 0.01)
```

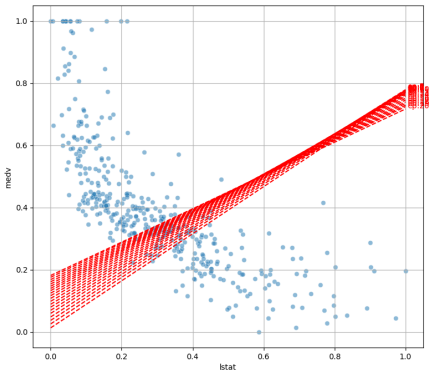


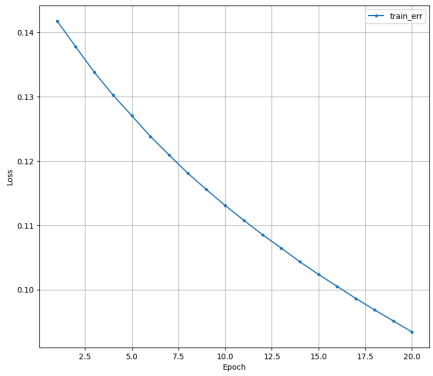


2) 실험2

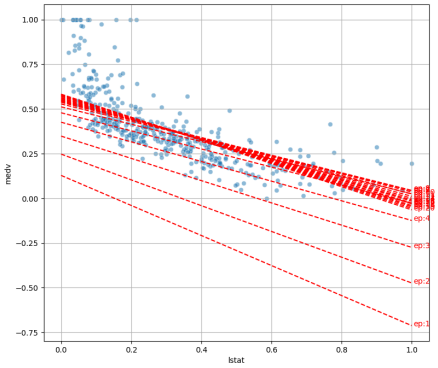
- epochs를 조절해 봅시다.
  - ep = 20으로 고정시켜 놓고
  - lr을 0.001, 0.01, 0.05, 0.1 로 조정하며 실행해 봅시다.

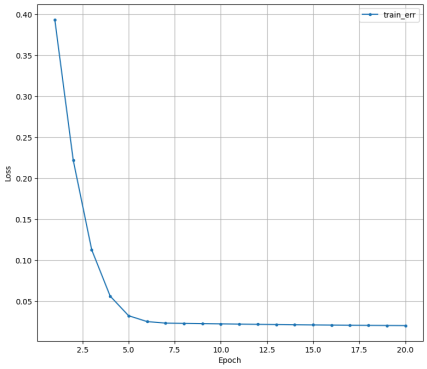
```
dl_visualize(ep = 20, lr = 0.001) # lr: lrearing_rate, 학습률, 가중치 조절의 크기에 영향을 미침
```



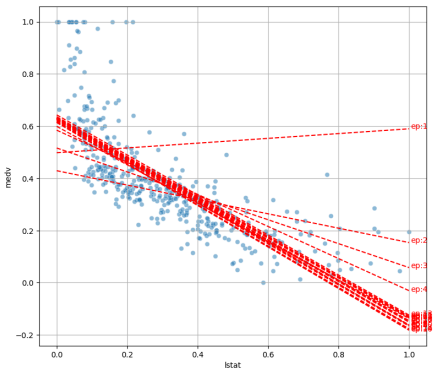


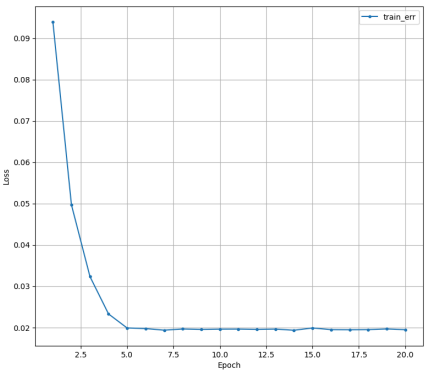

```
dl_visualize(ep = 20, lr = 0.01)
```



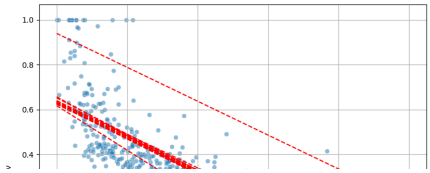


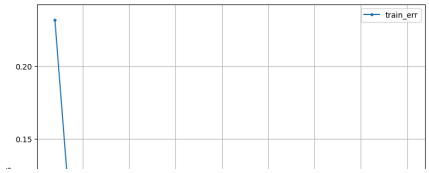

```
dl_visualize(ep = 20, lr = 0.05)
```






```
dl_visualize(ep = 20, lr = 0.1)
```


1.0
0.8
0.6
0.4


train_err
0.20
0.15