machine learning

1. Simple linear regression

```
pip install scikit-learn
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.10/dist-packages (1.2.2)
Requirement already satisfied: numpy>=1.17.3 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.22.4)
Requirement already satisfied: scipy>=1.3.2 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.10.1)
Requirement already satisfied: joblib>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (1.2.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from scikit-learn) (3.1.0)
```

step 1 import libraries

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

step 2 import data

```
df= pd.read_csv("salary_data.csv")
df.head()
```

	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2.0	43525
4	2.2	39891

step 3 selecting input and output variables

```
X=df[["YearsExperience"]]
y=df["Salary"]
```

```
X.head()
```

```
0    1.1
1    1.3
2    1.5
3    2.0
4    2.2
Name: YearsExperience, dtype: float64
```

```
y.head
```

```
<bound method NDFrame.head of 0      39343
1      46205
2      37731
3      43525
4      39891
5      56642
6      60150
7      54445
8      64445
9      57189
10     63218
11     55794
12     56957
13     57081
14     61111
```

```
15     67938
16     66029
17     83088
18     81363
19     93940
20     91738
21     98273
22    101302
23    113812
24    109431
25    105582
26    116969
27    112635
28    122391
29    121872
Name: Salary, dtype: int64>
```

step 4 making linear regression model

```
from sklearn.linear_model import LinearRegression
model=LinearRegression()
```

step 5 fitting a model

```
model= model.fit(X,y)
model
```

```
LinearRegression
LinearRegression()
```

step 6 predicting a model

```
model.predict([[50]])
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but LinearRegression was
  warnings.warn(
array([498290.31627142])
```

step 7 evaluating the model

```
from sklearn.model_selection import train_test_split
X_train , y_train , X_test,y_test=train_test_split(X,y,test_size=0.2,random_state=0)
# making and then fitting the model

model.fit(X_train,y_train)
```

✓ 0s completed at 10:29 AM