salary-prediction

October 22, 2024

[ ]:
```
import pandas as pd
import numpy as np
import plotly.express as px
import plotly.graph_objects as go

data = pd.read_csv("/content/Salary_Data.csv")
print(data.head())
```

```
   YearsExperience   Salary
0              1.1  39343.0
1              1.3  46205.0
2              1.5  37731.0
3              2.0  43525.0
4              2.2  39891.0
```

[ ]:
```
print(data.isnull().sum())
```

```
YearsExperience    0
Salary             0
dtype: int64
```

[ ]:
```
figure = px.scatter(data_frame = data,
                    x="Salary",
                    y="YearsExperience",
                    size="YearsExperience",
                    trendline="ols")
figure.show()
```

[ ]:
```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

x = np.asanyarray(data[["YearsExperience"]])
y = np.asanyarray(data[["Salary"]])
xtrain, xtest, ytrain, ytest = train_test_split(x, y,
                                                test_size=0.2,
                                                random_state=42)
```

[ ]:
```
model = LinearRegression()
model.fit(xtrain, ytrain)
```

[ ]:
```
LinearRegression()
```

[ ]:
```
a = float(input("Years of Experience : "))
features = np.array([[a]])
print("Predicted Salary = ", model.predict(features))
```

```
Years of Experience : 4
Predicted Salary =  [[63016.8443039]]
```