Machine learning

Multiple Linear Regression

Step 1.Import libraries

```
import pandas as pd
df = pd.read_csv("ml_data_salary.csv")
df.head()
```

	age	distance	YearsExperience	Salary
0	31.1	77.75	1.1	39343
1	31.3	78.25	1.3	46205
2	31.5	78.75	1.5	37731
3	32.0	80.00	2.0	43525
4	32.2	80.50	2.2	39891

Step 2. Import Dataset

```
X = df[["age","distance","YearsExperience"]]
y=df["Salary"]
```

Step 3. Fitting Linear Regression Model

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model = model.fit(X,y)
model
```

```
LinearRegression
LinearRegression()
```

Step 4.Evaluating the model fitness

```
# model fittness
print("Score for data =" , model.score(X,y))
```

```
Score for data = 0.9569960750337954
```

Step 5.Predicting the unknown vlues

```
model.predict([[31.1,77.75,1.1]])
```

```
/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([36209.375])
```

```
#####Step 6. check the accuracy according to 80/20
```

Step 6. check the accuracy according to 80/20

```
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import r2_score
model = LinearRegression()
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

```
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
accuracy = r2_score(y_test, y_pred)
print("Accuracy score: {:.2f}".format(accuracy))
```

```
Accuracy score: 0.99
```


✓ 0s completed at 8:04 AM