

Salary Prediction using Polynomial Regression

Importing the libraries

```
In [ ]: import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt
from sklearn.preprocessing import PolynomialFeatures
```

Loading the dataset

```
In [ ]: df=pd.read_csv('Salary.csv')
```

Summarize the Dataset

```
In [ ]: print(df.shape)
print(df.head())
```

```
(35, 2)
  YearsExperience  Salary
0              1.1   49343
1              1.3   56205
2              1.5   47731
3              2.0   53525
4              2.2   49891
```

```
In [ ]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35 entries, 0 to 34
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype  
---  -
0   YearsExperience  35 non-null    float64
1   Salary          35 non-null    int64   
dtypes: float64(1), int64(1)
memory usage: 692.0 bytes
```

```
In [ ]: df.describe()
```

Out[]:

	YearsExperience	Salary
count	35.000000	35.00000
mean	6.308571	93945.60000
std	3.618610	31719.74528
min	1.100000	47731.00000
25%	3.450000	67019.00000
50%	5.300000	93088.00000
75%	9.250000	123223.50000
max	13.500000	145675.00000

Segregate Dataset into Input X and Output Y

In []:

```
X=df.drop('Salary',axis='columns')
X
```

Out[]:

YearsExperience	
0	1.1
1	1.3
2	1.5
3	2.0
4	2.2
5	2.9
6	3.0
7	3.2
8	3.2
9	3.7
10	3.9
11	4.0
12	4.0
13	4.1
14	4.5
15	4.9
16	5.1
17	5.3
18	5.9
19	6.0
20	6.8
21	7.1
22	7.9
23	8.2
24	8.7
25	9.0
26	9.5
27	9.6
28	10.3
29	10.5
30	11.2
31	11.5
32	12.3

YearsExperience

33	12.9
34	13.5

```
In [ ]: Y=df.Salary
        Y
```

```
Out[ ]: 0      49343
        1      56205
        2      47731
        3      53525
        4      49891
        5      66642
        6      70150
        7      64445
        8      74445
        9      67189
       10      73218
       11      65794
       12      66957
       13      67081
       14      71111
       15      77938
       16      76029
       17      93088
       18     101363
       19     103940
       20     101738
       21     108273
       22     111302
       23     123812
       24     119431
       25     115582
       26     126969
       27     122635
       28     132391
       29     131872
       30     137345
       31     136756
       32     138765
       33     145675
       34     139465
        Name: Salary, dtype: int64
```

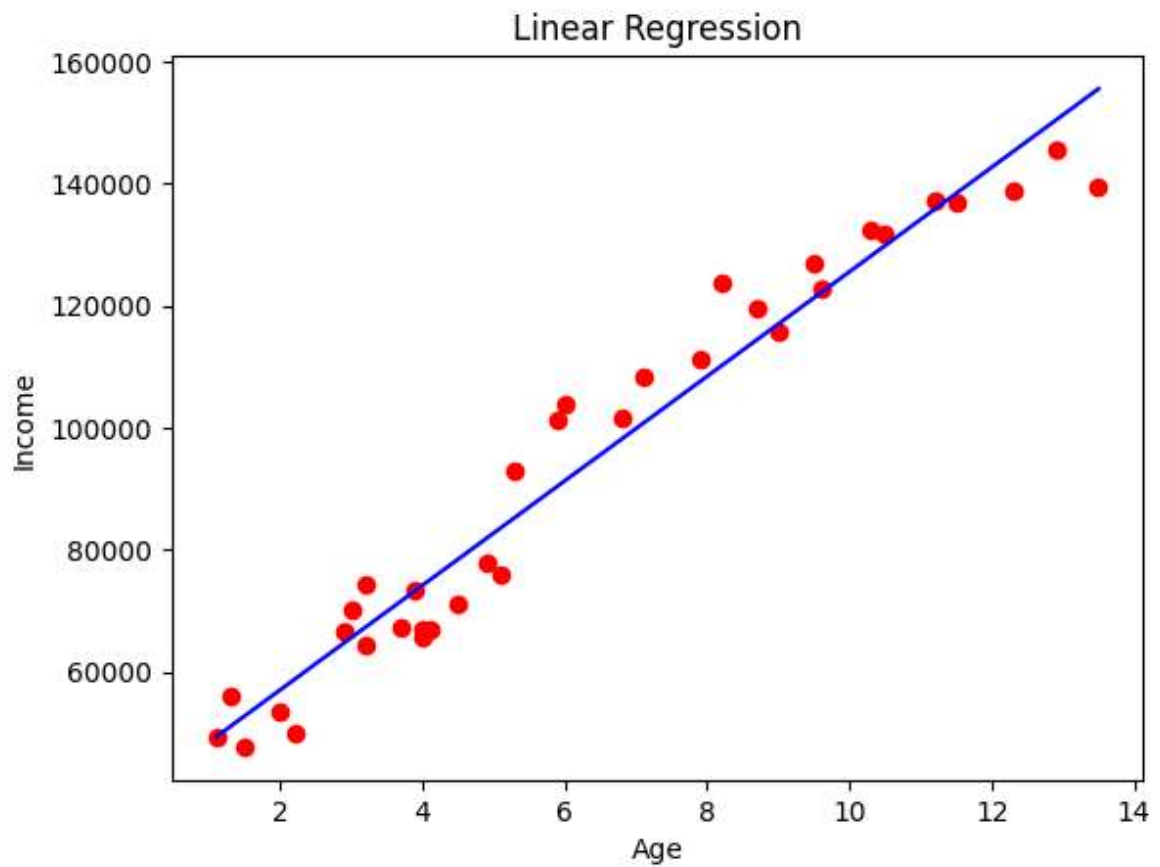
Training Dataset using Linear Regression

```
In [ ]: modelLR = LinearRegression()
        modelLR.fit(X,Y)
```

```
Out[ ]: ▾ LinearRegression
        LinearRegression()
```

Visualize the Linear Regression Results

```
In [ ]: plt.scatter(X,Y,color='red')
plt.plot(X,modelLR.predict(X),color='blue')
plt.title('Linear Regression')
plt.xlabel('Age')
plt.ylabel('Income')
plt.show()
```

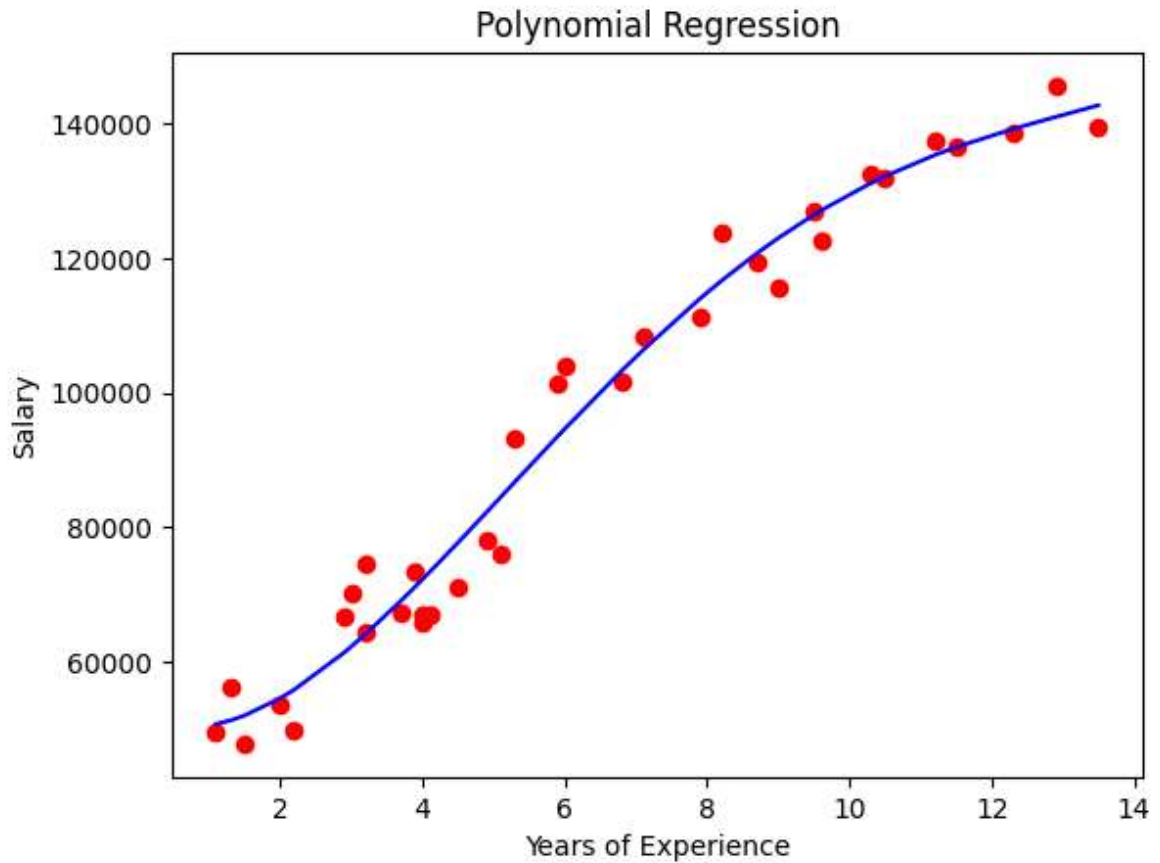


```
In [ ]: modelPR = PolynomialFeatures(degree=4)
X_poly = modelPR.fit_transform(X)
```

```
In [ ]: modelPLR = LinearRegression()
modelPLR.fit(X_poly,Y)
```

```
Out[ ]: ▾ LinearRegression
LinearRegression()
```

```
In [ ]: plt.scatter(X,Y,color='red')
plt.plot(X,modelPLR.predict(modelPR.fit_transform(X)),color='blue')
plt.title('Polynomial Regression')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.show()
```



```
In [ ]: x=float(input("Enter Years of Experience: "))
        salary = modelLR.predict([[x]])
        print("Salary of a person with age {} is {}".format(x,salary))
```

Salary of a person with age 8.5 is [112706.93301806]

C:\Users\pr12-\AppData\Local\Packages\PythonSoftwareFoundation.Python.3.11_qbz5n2kfra8p0\LocalCache\local-packages\Python311\site-packages\sklearn\base.py:464: UserWarning: X does not have valid feature names, but LinearRegression was fitted with feature names
warnings.warn(

```
In [ ]:
```