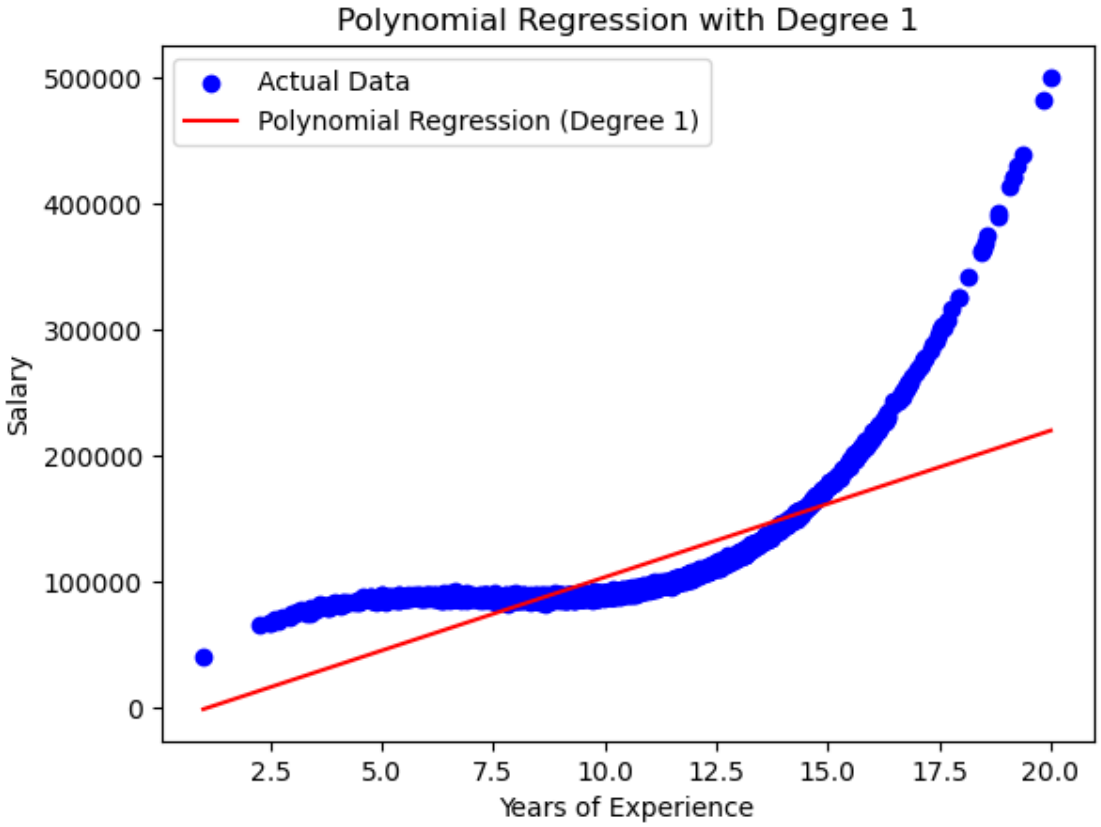
week5b

September 6, 2024

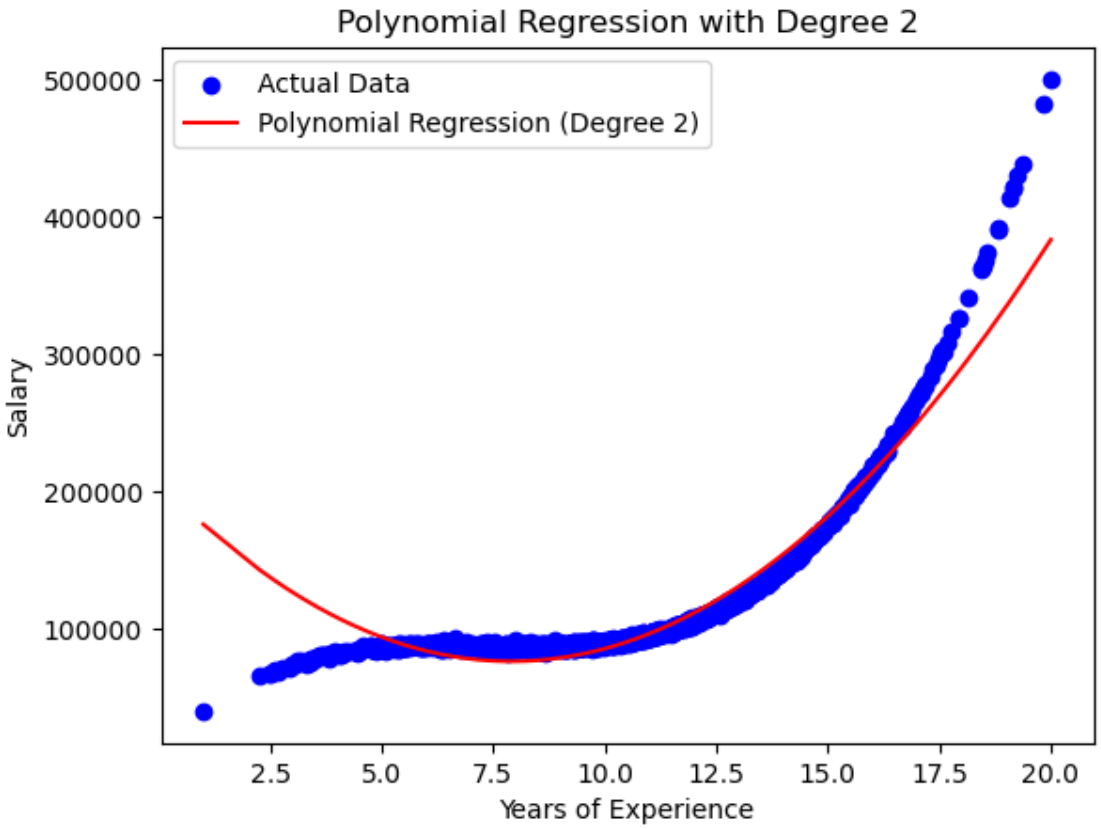
[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import PolynomialFeatures
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

df = pd.read_csv('employee_salary_dataset.csv')

X = df['Years of Experience'].values.reshape(-1, 1)
y = df['Salary'].values

degrees = [1, 2, 3, 4, 5]

for degree in degrees:

    poly_features = PolynomialFeatures(degree=degree)
    X_poly = poly_features.fit_transform(X)

    model = LinearRegression()
    model.fit(X_poly, y)

    y_pred = model.predict(X_poly)

    mse = mean_squared_error(y, y_pred)

    print(f'\nPolynomial Degree {degree}')
    print(f'Coefficients: {model.coef_}')
    print(f'Intercept: {model.intercept_}')
    print(f'Mean Squared Error (MSE): {mse}')

    plt.scatter(X, y, color='blue', label='Actual Data')
    plt.plot(X, y_pred, color='red', label=f'Polynomial Regression (Degree␣
 ↪{degree})')
    plt.title(f'Polynomial Regression with Degree {degree}')
    plt.xlabel('Years of Experience')
    plt.ylabel('Salary')
```

```
    plt.legend()
    plt.show()
```

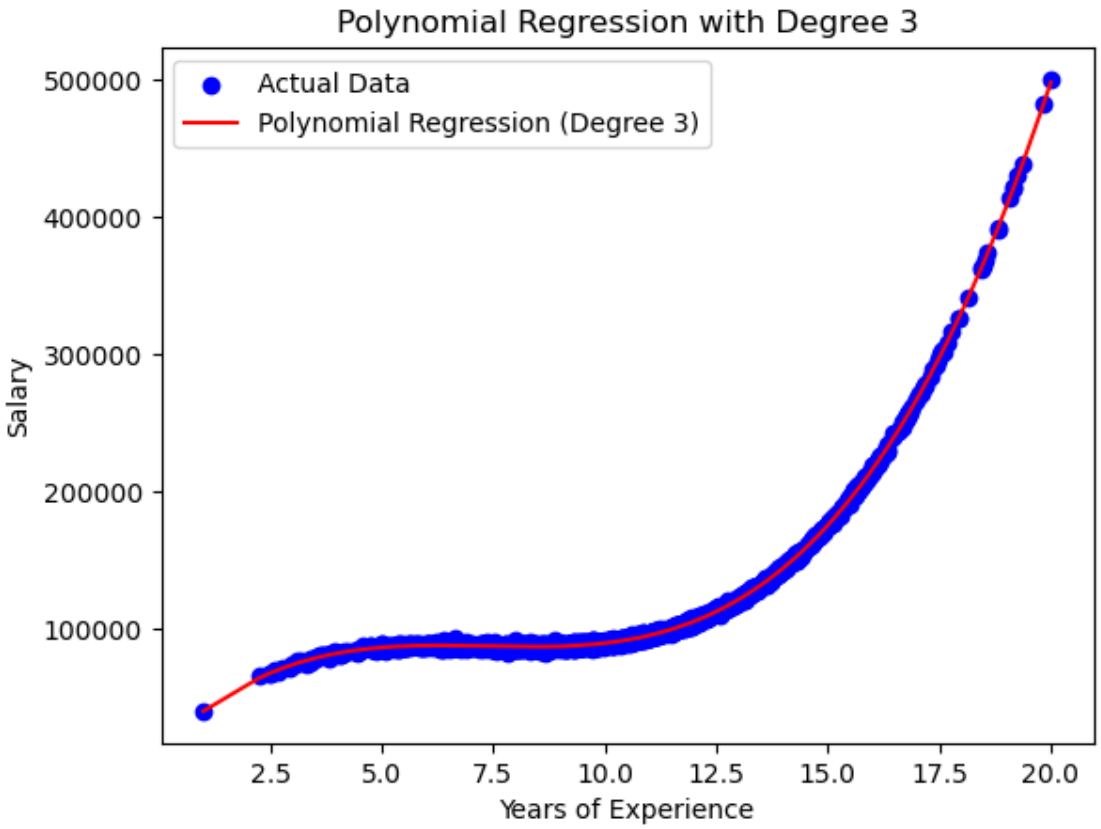
```
Polynomial Degree 1
Coefficients: [    0.         11640.22916165]
Intercept: -13063.099315186802
Mean Squared Error (MSE): 818398373.423397
```



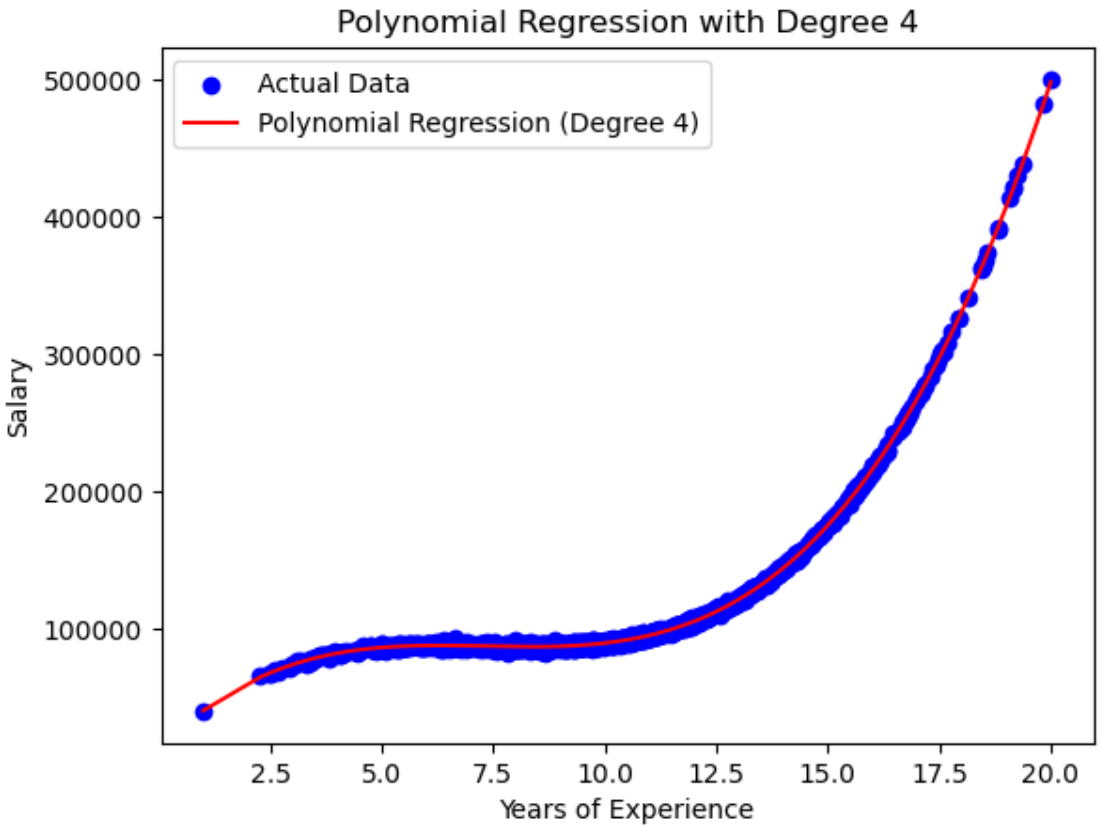

```
Polynomial Degree 2
Coefficients: [    0.         -33106.32895617   2095.78613853]
Intercept: 207045.95366531453
Mean Squared Error (MSE): 143553900.37966365
```



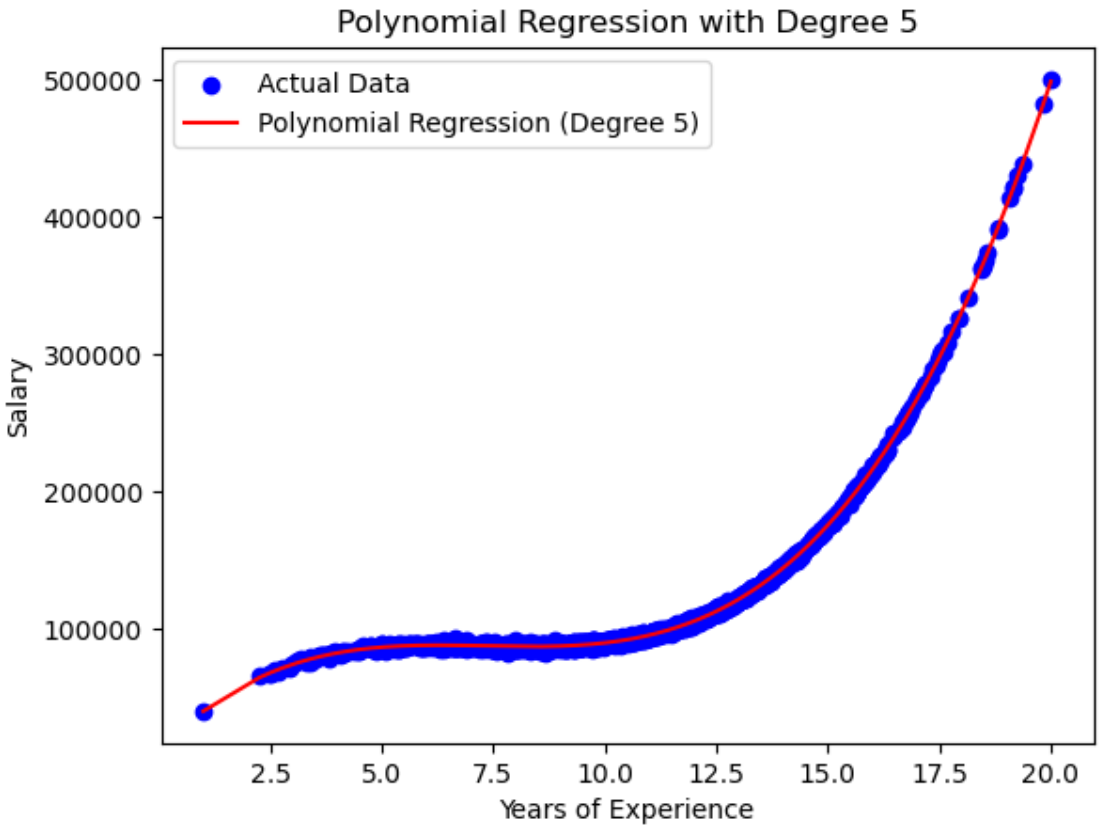

```
Polynomial Degree 3
Coefficients: [    0.         32261.24327461 -4503.00578987   205.20736238]
Intercept: 12053.107408621552
Mean Squared Error (MSE): 1627265.1068829503
```




```
Polynomial Degree 4
Coefficients: [ 0.00000000e+00  3.18998938e+04 -4.44416359e+03  2.01309144e+02
  9.01918620e-02]
Intercept: 12793.58737821845
Mean Squared Error (MSE): 1626600.7213873735
```




```
Polynomial Degree 5
Coefficients: [ 0.00000000e+00  3.27949771e+04 -4.65578675e+03  2.24001236e+02
 -1.03521070e+00  2.09284278e-02]
Intercept: 11462.338086193631
Mean Squared Error (MSE): 1625733.1261048843
```



- The code fits a polynomial regression model for different degrees (1 to 5) to predict salary based on years of experience.
- For each degree, it transforms the feature data, trains a LinearRegression model, computes the coefficients, intercept, and Mean Squared Error (MSE), then prints these results.
- It plots the actual data points against the predicted values, showing how the polynomial model fits the data for each degree.