IML Assignment # 1

Report

Problem Statement:

I collected the data of different houses of Bahria Town Islamabad from zameen.com website. The data having features like location, area, number of beds, number of baths and price. I trained the model using Multivariate and Polynomial regression to predict the price of the houses. Here are the results that I get after training the model on the data set that I collected.

Results:

Multivariate Regression

(Without Feature Scaling)

```
Mean Squared Error:  4.831565660405444

Learning Rate:  0.001
Mean Squared Error:  4.1945878714675885e+23

Learning Rate:  0.01
Mean Squared Error:  5.0047128559581983e+23

Learning Rate:  0.1
Mean Squared Error:  1.66508475394773e+27

Learning Rate:  0.5
Mean Squared Error:  3.8244616152954563e+28

Learning Rate:  1
Mean Squared Error:  1.7158820568654663e+29
```

(With Feature Scaling)

There was no need of feature scaling because negative values obtained after scaling.

```
Learning Rate:  0.001
Mean Squared Error:  5.587585095665656

Learning Rate:  0.01
Mean Squared Error:  5.5777547169955195

Learning Rate:  0.1
Mean Squared Error:  5.10727109770817

Learning Rate:  0.5
Mean Squared Error:  41.02136007653824

Learning Rate:  1
Mean Squared Error:  3.831660528847993e+22
```

Polynomial Regression

(Without Feature Scaling)

```
Mean Squared Error Polynomial Regression:
4.78528733436451


Learning Rate:  0.001

Mean Squared Error:  4.9165684234088947e+33


Learning Rate:  0.01

Mean Squared Error:  4.2911137288906214e+35
```

```
Learning Rate:  0.1
Mean Squared Error:  4.2944227694358 38e+37

Learning Rate:  0.5
Mean Squared Error:  1.0738738365048 78e+39

Learning Rate:  1
Mean Squared Error:  4.2611184285764 38e+39
```

(With Feature Scaling)

There was no need of feature scaling because negative values obtained after scaling.

```
Learning Rate:  0.001
Mean Squared Error:  4.901112256067842

Learning Rate:  0.01
Mean Squared Error:  4.785950157912555

Learning Rate:  0.1
Mean Squared Error:  5.0221816695956765

Learning Rate:  0.5
Mean Squared Error:  3.8352821545616684e+22

Learning Rate:  1
Mean Squared Error:  9.579838193932118e+23
```

Normal Equation:

The score I obtained using normal equation was:

```
Score of Normal Equation: 4.831565660405443
```

Conclusion:

In polynomial Regression I am getting minimum mean squared error of 1.07 at learning rate of 0.5 (without feature scaling). While larger errors like 5.00, 4.87…etc and 4.83 of normal equation. This may be due to overfitting and underfitting and there is no need of feature scaling because of getting negative values after feature scaling. It's possible that this model is performing well on the training set, but it may not generalize well to new, unseen data.