# Stochastic Gradient Descent

Machine Learning Exercise 1

May 3, 2020

Alina Munir (7251819)  Anne kunstmann (6651481)

Submitted to: DR. VICTOR UC CETINA

# Stochastic Gradient Descent

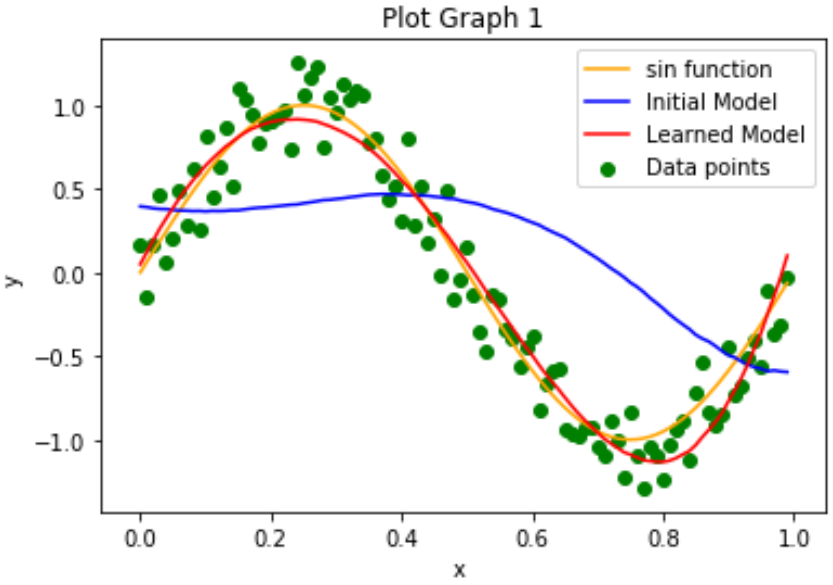
1) Different degrees of polynomials were tested to generate a model for linear regression of 100 (x,y) points. With an increase in degree D, more desired behaviour in the model was found and at **D=4** model adopted its ultimate position of improvement but for D>4 model started to adapt to sine function curve more than the scattered points on the graph, which was not the goal of our regression. Hence **polynomial of degree 4** was chosen as the model.

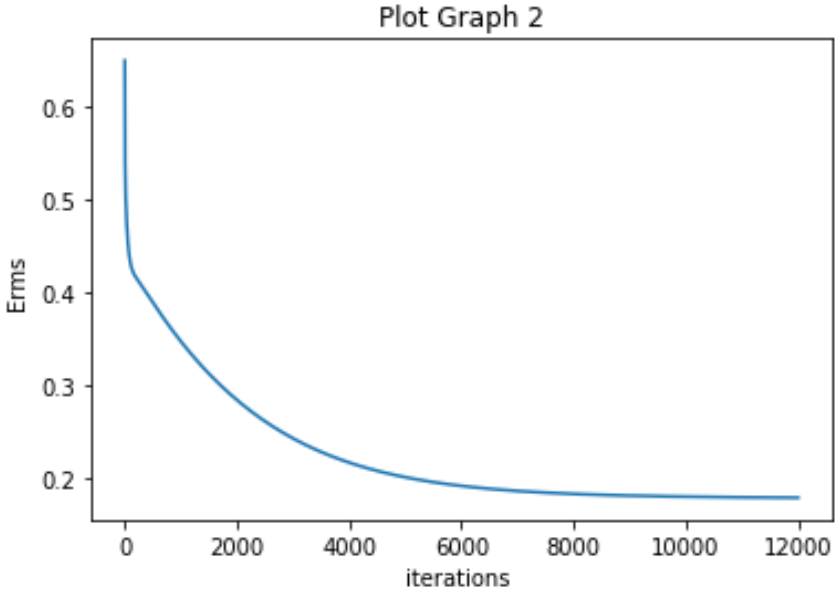
   Initial values of θi parameters were randomly generated in the interval [−0.5,0.5]. Over **12000 iterations** we have managed to minimize the error through Stochastic Gradient Descent. Model's initial and final parameters can be distinguished from following unlearnt and learnt model polynomials:

```
Initial Model h0(x) =  0.3948775534013792  +  -0.1842739864158297 x +
0.13125736791029274 x^2 +  -0.3538250055920371 x^3 0.11438900413773123 x^4

Final Model h0(x) =  0.04815698592377191  +  7.677207165333941 x +
-16.6754070534138 x^2 +  -3.8811686965687606 x^3 13.076441136538023 x^4
```

2) Different alpha values were tested for the polynomial but **α = 0.01** has been chosen due to following results:
   - **a)** For α<0.01 achieved model was not even close to our desired model and error didn't stable even after 30000 iterations.
   - **b)** For α=0.01 model adapted its optimal curve just after 12000 iterations and error value got greatly stabilized.
   - **c)** For α>0.01 model curve started inclining towards every small cluster of points, hence created glitches in the curve like a noised/distorted sine function. Although, our model learnt very fast in a perfect slope just under 3000 iterations and then stabilize


Plot Graph 1
sin function
Initial Model
Learned Model
Data points
1.0
0.5
0.0
−0.5
−1.0
y
0.0
0.2
0.4
0.6
0.8
1.0
x


Plot Graph 2
0.6
0.5
0.4
0.3
0.2
Erms
0
2000
4000
6000
8000
10000
12000
iterations