Long Nguyen Hoang
CS550

Week 2: Homework 1: Using Overfitting to evaluate Models

X	y	y=a1+b1*x	y=a2+b2*x^2	x	y	y=a1+b1*x	y=a2+b2*x^2	x	y=a1+b1*x or y=a2+b2*x^2
1	1.8	1.38	1.8	1.5	1.7	1.81	1.9625	1.4	1.724
2	2.4	2.24	2.19	2.9	2.7	3.014	2.7633	2.5	2.4825
3.3	2.3	3.358	3.0857	3.7	2.5	3.702	2.4825	3.6	3.3548
4.3	3.8	4.218	4.0737	4.7	2.8	4.562	2.6892	4.5	4.3025
5.3	5.3	5.078	5.3217	5.1	5.5	4.906	5.6025	5.4	5.164
1.4	1.5	1.724	1.9248	x	x	x	x	x	x
2.5	2.2	2.67	2.4825	x	x	x	x	x	x
2.8	3.8	2.928	2.6892	x	x	x	x	x	x
4.1	4.0	4.046	3.8553	x	x	x	x	x	x
5.1	5.4	4.906	5.0513	x	x	x	x	x	x

Training phase N=10

X	y	x*y	x*x
1	1.8	1.8	1
2	2.4	4.8	4
3.3	2.3	7.59	10.89
4.3	3.8	16.34	18.49
5.3	5.3	28.09	28.09
1.4	1.5	2.1	1.96
2.5	2.2	5.5	6.25
2.8	3.8	10.64	7.84
4.1	4.0	16.4	16.81
5.1	5.4	27.54	26.01

SUM x = 31.8
SUM y = 32.5
SUM x*y= 120.8
SUM x*x= 121.34

Slope(b) = (N*SUMx*y - SUMx*SUMy) / (N*SUMx*x - SUMx^2)
Slope(10) = (10*120.8 - 31.8*32.5) / 10*121.34 - 31.8^2=
=(1208 - 1033.5) / (1213.4 - 1011.24)=
=174.5/202.16=0.86

Intercept(a) = (SUMy - b(SUMx))/N=
=(32.5 - 0.86*31.8) / 10=
=(32.5 - 27.348)/10=
=5.152/10=0.52

Non-linear

X	x*x	y
1	1	1.8
2	4	2.4
3.3	10.89	2.3
4.3	18.49	3.8
5.3	28.09	5.3
1.4	1.96	1.5
2.5	6.25	2.2
2.8	7.84	3.8
4.1	16.81	4.0
5.1	26.01	5.4

x*x	y	x^2*y	x^2*x^2
1	1.8	1.8	1
4	2.4	9.6	16
10.89	2.3	25.047	118.25921
18.49	3.8	70.262	341.8801
28.09	5.3	148.877	789.0481
1.96	1.5	2.94	3.8416
6.25	2.2	13.75	39.0625
7.84	3.8	29.792	61.4656
16.81	4.0	67.24	282.5761
26.01	5.4	140.454	676.5201

SUM x*x= 121.34
SUM y = 32.5
SUM x^2*y = 509.762
SUM x^2*x^2 = 2329.65331

Slope(b) = (N*SUMx^2*y - SUMx*xSUMy) / (N*SUMx^2*x^2 - (SUMx*x)^2) =
=(10*509.726) - 121.34*32.5 / 10*2329.65331 - 121.34^2 =
=5097.26 - 3943.55 / 23296.5331 - 14723.3956=
=1153.71 / 8573.1375 = 0.13

Intercept(a) = (SUMy - b(SUMx*x)) / N =
=(32.5 - 0.13*121.34) / 10=
=(32.5 - 15.7742) / 10 =
= 16.7258 / 10 = 1.67