

Id	Sq. Ft	Lot	Beds	Baths	Year	Price	Prediction W=[10,1,1,1,1] b=10,000	MSE
1	1826	19378	4	2.5	2005	320000	49649.5	5.22778E+11
2	1050	7500	2	2	2004	185000	30008	
3	1276	12209	3	2.5	2001	215000	36975.5	
4	1040	7658	2	2	2005	319900	30067	
5	1535	4500	2	2	1998	164000	31852	
6	1535	13704	3	2	2001	205000	41060	
7	1040	11143	4	3	2004	340000	33554	
8	1370	13005	4	2.5	1980	260000	38691.5	
9	2036	10207	3	3	2007	227875	42580	
10	2899	13682	3	3.5	2006	438780	54684.5	
Part 3: The reason why squared in necessary in evaluating the performance of a model is to measure the distance from the actual value, so if its negative, it is irrelevant. Another way to do this is to take the absolute value, which would make all the numbers positive					Part 4: I believe the best training model is #1, it did not overshoot the lowest mean square error, and was more efficient in reaching that answer than training #2, training number 3 did not seem to get any better and oscillated between 2 MSE values			
Part 5: The reason that the model is perfectly accurate on only the provided data set is because the model is overfit for the training data, and when presented with new data it is too specialized to be correct in its prediction.					Part 6: In order to fix the overfitting problem, we can employ a method discussed in class, Early stopping. If we stop the algorithm before it makes perfectly tailored answers to that one data set, it will be more generalized and be able to apply its knowledge to different sets of data.			