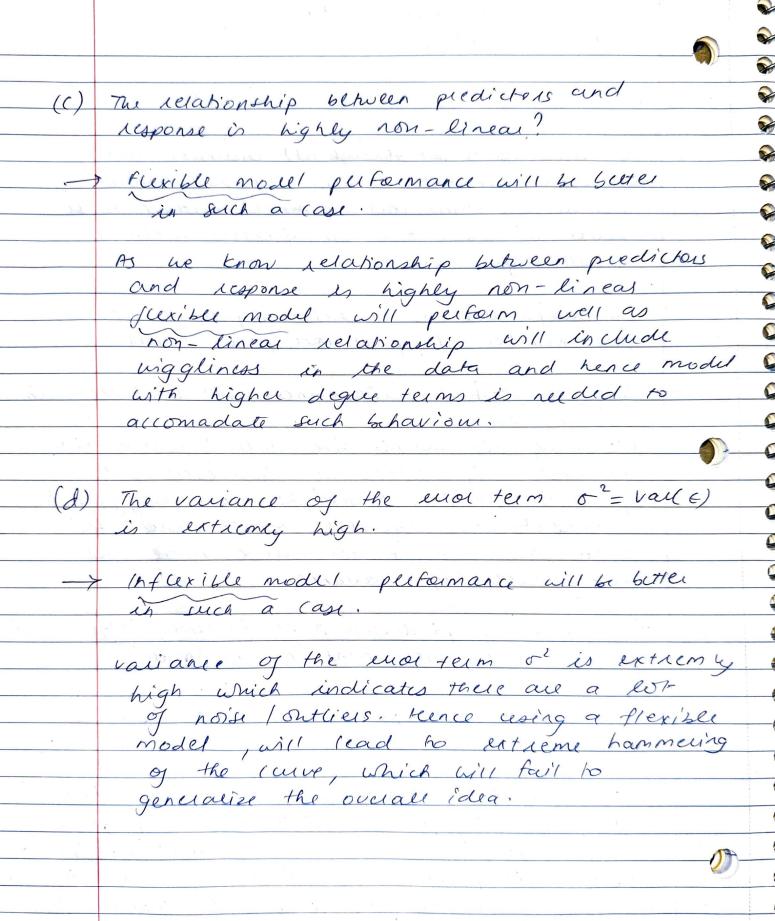
ANISH RAJESH ADNANI USCED: 4092610491 USC Email: adnani@usc.edu githes: midnightbox for each parts (a) through (d) indicate whether we would generally expect the performance of a flexible statistical learning model method to be better or worse than an flexible method. Justify your arswer The sample size n is extremly large, and the number of predictors p is small. Flexible Method performance will be better in such a case. Increasing the sample size will increase the variance, if variance is increased the plediction model should be more wiggly to accomadate herce furible model should be used to hamme the cure more to accompance high variance (b) The number of predictors p is extremely large and the number of observations 'n' is small? such a case, 3 with small dataset and large predictors, there is high probability that flexible model will overfit as it will try to hammer the cure to fit to all the datapoints with 'n' being small there is a high chance flexible model becomes like a lookup table, heave failing to generalise the



ISL R 2.7 The table selow provides a training data set Containing six observations, there predictors, and one ovalitative response variable. 065 12 X3 Red Red Red 3 3 green green 1 Red. suppose we wish to use this data set to make a predictor for Y when x1 = x2 = x3 = 0 using K-nearest reighbors. compute the tuclidean distance between each (a) observation and test point x1=0 x2 =0 x3=6 Euclidean dist = \(\tau_1 - \tau_2 \tau_1 + (4, - 42)^2 + \dots dist 065 V(0-0) + (0-3) + (0-0) =3 V (2-0) + (0-0) + (0-0)2 = 2 $\sqrt{(0-0)^2 + (1-0)^2 + (3-0)^2 - 3.1622}$ 1/10-0)2+(1-0)2+(2-0)2=2-2360 $\sqrt{(-1-0)^2 + (0-0)^2 + (1-0)^2} = 1.414$ N (1-0)2+ (1-0)2 + (1-0)2 = 1.732

(6) what is our prediction with K=12 Why? -> MIN(3,2,3.1622, 2.2360, 1.414, 1.732 with K=1 the given point is nearest to test dara point @ which is (-1,0,1) which is quen herce with K=1 given point (0,0,0) will ask be assigned (() What is one prediction of K=3? why? 3,2,3.6622, 2.2360, 1.414, 1.732 1.414, 1.732, 2, 2.2360, 3, 3.1622 (D (D) dist nearest 3 point ares (-1, 0, 1) green Red Red. (2,0,0) Since majority of these 3 labels is Red red is assigned to given data posit (0,9,9) (d) If the Bayes decision coundary in this
problem is highly non-linear then would
we expect the best value of the be
large or small? Why? Bayes decision boundary is highly non-linear which indicates there is high variance in data, which means we need a wiggly fit we need a smaller value of K to make the curve male wiggly rather than linear as in raige value