

Chapter 7

If predictors are very different in scales, should we standardize predictors before running kNN?

Yes

Consider a kNN model with $k=7$. For a new observation, there are 5 nearest neighbors in class 1 and 2 nearest neighbors in class 0. With a cutoff = 0.75, which class should we assign this observation to?

Class 0. $C_1 = 5$, $C_0 = 2$. $C_1 = (5/7 = 0.71)$. $C_2 = (2/7 = 0.28)$. $Y=0 \rightarrow (p(0.71) < \text{cutoff}(0.75)) \rightarrow \text{Class 0}$

Which of the following weights are reasonable for kNN of prediction? Denote by d_i the distance of the i -th nearest neighbor.

$(1/d_i) / (1/d_1 + 1/d_2 + \dots + 1/d_k)$

Consider the same 2-class classification problem and use the same output table as in the previous question. Which value for k is most appropriate?

3

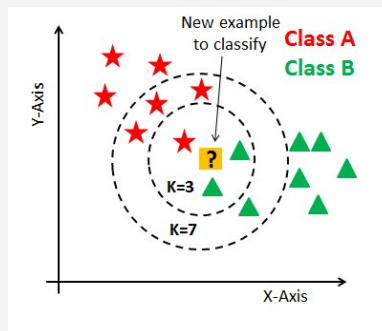
Validation Error with Different K Values

	A	B	C	D	E	F
50	Validation error log for different k					
51						
52						
53						
54						
55						
56						
57						
58						
59						
60						
61						
62						
63						

Continuous: You and XLMiners will pick the LOWEST one.

Binary: Pick the lowest non-even K

Probability new observation belongs to class A



In kNN with $k = 3$, what is the probability that the new observation belongs to class A?

Class A: C_0 : 1

Class B: C_1 : 2

$K = 1/3 \rightarrow \# \text{ in class A over } K = \#$

Chapter 9

For continuous variable with 100 unique values, how many possible partitions are there along that variable in the recursive partitioning

99

Suppose we are studying a 4-class classification problem with classification tree.

What is the maximum value of entropy measure of impurity?

$2 \rightarrow \log_2(4) = 2$

How many possible partitions can we have for a categorical predictor with 4 categories? (Hint: you can enumerate them)

$7 \rightarrow abcd, a-bcd, b-acd, c-dab, d-abc, ab-cd, ac-bd, ad-ac$

Is a full tree over fitted

Yes

Which tree is smaller, minimum error tree or best pruned tree? Suppose they are different (sometimes they can be the same tree)

Best Pruned Tree

Chapter 9 (cont)

How do we measure the impurity of a partition in regression tree?

Sum of Squared Deviations

Is CART a model free algorithm

Yes

chapter 10

For a logistic regression model estimated as below, what is the probability of accepting personal loan offer for a person with income of 50K dollars ($X = 50$)?

1.22% i think. use the regression equation and sub in for x. $= (1 / (1 + (\exp(-6.3525 - 0.0392 * (50))))) = \text{probability}$

Using the cutoff of 10%, should this person be classified as who will accept the offer?

Yes

What are the Odds of accepting the loan offer for this person?

$0.0124 \rightarrow p / (1 - P) = \text{the odds}$

What is the odds ratio of income?

1.03998

Is logistic regression a model-free algorithm

NO