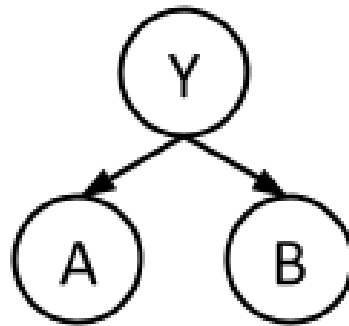


# 1 Naive Bayes

In this question, we will train a Naive Bayes classifier to predict class labels  $Y$  as a function of input features  $A$  and  $B$ .  $Y$ ,  $A$ , and  $B$  are all binary variables, with domains 0 and 1. We are given 10 training points from which we will estimate our distribution.

$A$	1	1	1	1	0	1	0	1	1	1
$B$	1	0	0	1	1	1	1	0	1	1
$Y$	1	1	0	0	0	1	1	0	0	0



(a) What are the maximum likelihood estimates for the tables  $P(Y)$ ,  $P(A|Y)$ , and  $P(B|Y)$ ?

$Y$	$P(Y)$
0	
1	

$A$	$Y$	$P(A Y)$
0	0	
1	0	
0	1	
1	1	

$B$	$Y$	$P(B Y)$
0	0	
1	0	
0	1	
1	1	

(b) Consider a new data point ( $A = 1, B = 1$ ). What label would this classifier assign to this sample?

(c) Let's use Laplace Smoothing to smooth out our distribution. Compute the new distribution for  $P(A|Y)$  given Laplace Smoothing with  $k = 2$ .

$A$	$Y$	$P(A Y)$
0	0	
1	0	
0	1	
1	1	

## 2 Perceptron

You want to predict if movies will be profitable based on their screenplays. You hire two critics A and B to read a script you have and rate it on a scale of 1 to 4. The critics are not perfect; here are five data points including the critics' scores and the performance of the movie:

#	Movie Name	A	B	Profit?
1	Pellet Power	1	1	-
2	Ghosts!	3	2	+
3	Pac is Bac	2	4	+
4	Not a Pizza	3	4	+
5	Endless Maze	2	3	-

- (a) First, you would like to examine the linear separability of the data. Plot the data on the 2D plane above; label profitable movies with + and non-profitable movies with - and determine if the data are linearly separable.
- (b) Now you decide to use a perceptron to classify your data. Suppose you directly use the scores given above as features, together with a bias feature. That is  $f_0 = 1$ ,  $f_1 =$  score given by A and  $f_2 =$  score given by B.

Run one pass through the data with the perceptron algorithm, filling out the table below. Go through the data points in order, e.g. using data point #1 at step 1.

step	Weights	Score	Correct?
1	$[-1, 0, 0]$	$-1 \cdot 1 + 0 \cdot 1 + 0 \cdot 1 = -1$	yes
2			
3			
4			
5			

Final weights:

- (c) Have weights been learned that separate the data?
- (d) More generally, irrespective of the training data, you want to know if your features are powerful enough to allow you to handle a range of scenarios. Circle the scenarios for which a perceptron using the features above can indeed perfectly classify movies which are profitable according to the given rules:
- Your reviewers are awesome: if the total of their scores is more than 8, then the movie will definitely be profitable, and otherwise it won't be.
  - Your reviewers are art critics. Your movie will be profitable if and only if each reviewer gives either a score of 2 or a score of 3.
  - Your reviewers have weird but different tastes. Your movie will be profitable if and only if both reviewers agree.