Exercise 15 : Probability, Bayes Optimal Classifier, Decision Tree

Exercise 15-1 : Probability

The number of notebooks sold in a small computer store during a week is described by a random variable X having probability function

$$f(x) = \frac{2x+3}{63}, \quad x \in R_X = \{0, 1, 2, 3, 4, 5, 6\}.$$

(a) Find the expected number of notebooks sold by the store in a given one-week period.

(b) In order to have enough stock, the store orders every week from its supplier 6 notebooks at a price of \$250 each, under the following agreement : the new notebooks arrive at the store on Monday morning and any notebook not sold during the week can be returned to the supplier at the price of \$210. If the store sells notebooks at a price of \$325, find the store's expected profit during a week.

Exercise 15-2 : Bayes Optimal Classifier

We have a classification problem with three classes $1, 2, 3$, and three trained classifiers h_1, h_2, and h_3, with the following probabilities of the classifiers, given the training data D :

$$\Pr(h_1|D) = 0.25$$
$$\Pr(h_2|D) = 0.35$$
$$\Pr(h_3|D) = 0.4$$

For the four test instances o_1, o_2, o_3, o_4, the classifiers give the following class probabilities :

o_1	$\Pr(1\|h_1) = 0.3$	$\Pr(2\|h_1) = 0.4$	$\Pr(3\|h_1) = 0.3$
	$\Pr(1\|h_2) = 0.25$	$\Pr(2\|h_2) = 0.55$	$\Pr(3\|h_2) = 0.2$
	$\Pr(1\|h_3) = 0.15$	$\Pr(2\|h_3) = 0.6$	$\Pr(3\|h_3) = 0.25$
o_2	$\Pr(1\|h_1) = 0.4$	$\Pr(2\|h_1) = 0.3$	$\Pr(3\|h_1) = 0.3$
	$\Pr(1\|h_2) = 0.45$	$\Pr(2\|h_2) = 0.25$	$\Pr(3\|h_2) = 0.3$
	$\Pr(1\|h_3) = 0.55$	$\Pr(2\|h_3) = 0.1$	$\Pr(3\|h_3) = 0.35$
o_3	$\Pr(1\|h_1) = 0.4$	$\Pr(2\|h_1) = 0.5$	$\Pr(3\|h_1) = 0.1$
	$\Pr(1\|h_2) = 0.5$	$\Pr(2\|h_2) = 0.4$	$\Pr(3\|h_2) = 0.1$
	$\Pr(1\|h_3) = 0.15$	$\Pr(2\|h_3) = 0.1$	$\Pr(3\|h_3) = 0.75$
o_4	$\Pr(1\|h_1) = 0.35$	$\Pr(2\|h_1) = 0.5$	$\Pr(3\|h_1) = 0.15$
	$\Pr(1\|h_2) = 0.2$	$\Pr(2\|h_2) = 0.5$	$\Pr(3\|h_2) = 0.3$
	$\Pr(1\|h_3) = 0.15$	$\Pr(2\|h_3) = 0.7$	$\Pr(3\|h_3) = 0.15$

We combine the three classifiers to get a Bayes optimal classifier. Which class probabilities will we get from this Bayes optimal classifier for the four test instances ?

Exercise 15-3 : Decision Tree

Given the training data about trading in the table below :

RID	Past Trend	Open Interest	Trading Volume	Return
1	Positive	Low	High	Up
2	Negative	high	Low	Down
3	Positive	Low	High	Up
4	Positive	High	High	Up
5	Negative	Low	High	Down
6	Positive	Low	Low	Down
7	Negative	High	High	Down
8	Negative	Low	High	Down
9	Positive	Low	Low	Down
10	Positive	High	High	Up

(a) Create a decision tree based on Gini index.

(b) Classify the test instances :

i) Past Trend = Positive, Open Interest = High, Trading Volume = High

ii) Past Trend = Negative, Open Interest = Low, Trading Volume = High