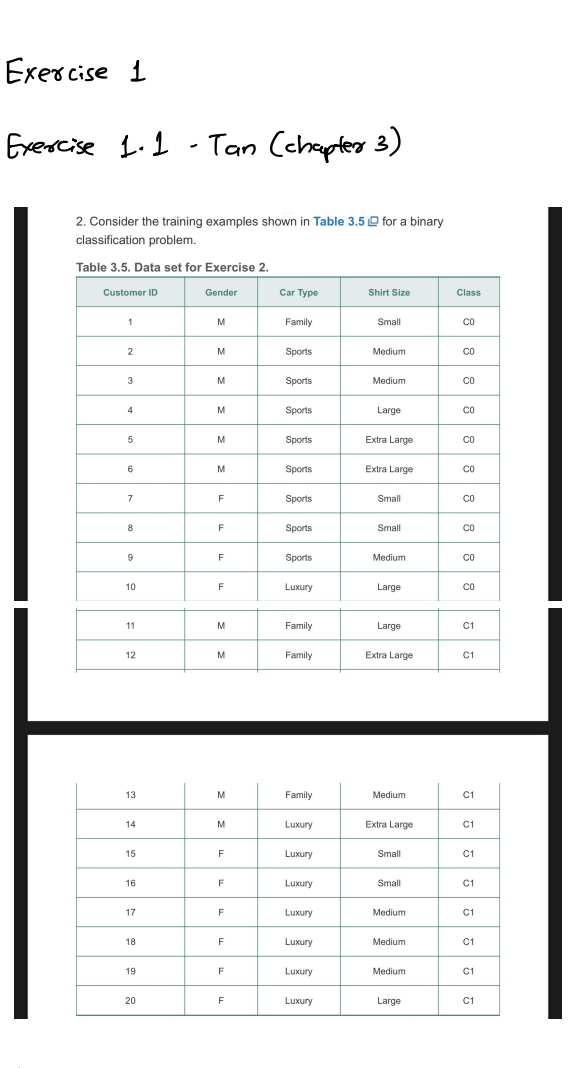
CS 422 HW4 (A20516822)

Wednesday, 5 October 2022 12:31 PM

Exercise 1

Exercise 1.1 - Tan (chapter 3)



2. Consider the training examples shown in Table 3.5 for a binary classification problem.

Table 3.5. Data set for Exercise 2.

Customer ID	Gender	Car Type	Shirt Size	Class
1	M	Family	Small	C0
2	M	Sports	Medium	C0
3	M	Sports	Medium	C0
4	M	Sports	Large	C0
5	M	Sports	Extra Large	C0
6	M	Sports	Extra Large	C0
7	F	Sports	Small	C0
8	F	Sports	Small	C0
9	F	Sports	Medium	C0
10	F	Luxury	Large	C0
11	M	Family	Large	C1
12	M	Family	Extra Large	C1
13	M	Family	Medium	C1
14	M	Luxury	Extra Large	C1
15	F	Luxury	Small	C1
16	F	Luxury	Small	C1
17	F	Luxury	Medium	C1
18	F	Luxury	Medium	C1
19	F	Luxury	Medium	C1
20	F	Luxury	Large	C1

(a) Compute the Gini index for the overall collection of training examples.

Answer:-

Gini Index is given as $1-\sum_{i=0}^{c-1} p_i(t)^2$

As we have 10 instances of class C0 & 10 instances of class C1, Hence no. of classes, c = 2

$$\text{Gini Index} = 1-\sum_{i=0}^{1} p_i(t)^2$$
$$= 1-\left[p_0(t)^2 + p_1(t)^2\right]$$
$$= 1-\left[\left(\frac{10}{20}\right)^2+\left(\frac{10}{20}\right)^2\right]$$
$$= 1-\left[2\times(0.5)^2\right]$$
$$= 1-0.5$$
$$= 0.5$$

(b) Compute the Gini index for the Customer ID attribute.

Answer:-

For each instance, we have a unique customer ID value

Hence it will be a N-split case. As we have 20 instances, it will be a 20 way split.

Gini Index for each is given by

$$G.I = 1-\left[\left(\frac{1}{1}\right)^2\right] = 1-1 = 0$$

$$\text{Weighted Gini Index for Customer Id} = \left[\frac{1}{20}\times 0\right]\times 20$$
$$= 0$$

(c) Compute the Gini Index for Gender attribute.

Answer:-

For this, we have two possibilities, Gender = Male & Gender = female

For Gender = Male, 6 instances belong to C0, 4 instances belong to C1

$$\text{Gini Index (Gender = Male)} = 1-\left[\left(\frac{6}{10}\right)^2+\left(\frac{4}{10}\right)^2\right]$$
$$= 1-\left[(0.36)+(0.16)\right]$$
$$= 1-0.52$$
$$= 0.48$$

For Gender = Female, 4 instances belong to C0, 6 instances belong to C1

$$\text{Gini Index (Gender = female)} = 1-\left[\left(\frac{4}{10}\right)^2+\left(\frac{6}{10}\right)^2\right]$$
$$= 1-\left[0.16+0.36\right]$$
$$= 1-0.52$$
$$= 0.48$$

$$\text{Weighted Gini Index (Gender)} = \frac{10}{20}\times 0.48 + \frac{10}{20}\times 0.48$$
$$= 0.48$$

(d) Compute the Gini Index for Car Type attribute using multiway split.

Answer:-

Using a multiway split, we have 3 possible values for CarType

For CarType = Family, 1 instance belong to C0, 3 instances belong to C1.

$$\text{Gini Index (CarType = Family)} = 1-\left[\left(\frac{1}{4}\right)^2+\left(\frac{3}{4}\right)^2\right]$$
$$= 1-\left[(0.25)^2+(0.75)^2\right]$$
$$= 1-\left[0.0625+0.5625\right]$$
$$= 1-0.625$$
$$= 0.375$$

For CarType = Sports, 8 instances belong to C0, 0 instances belong to C1

$$\text{Gini Index (CarType = Sports)} = 1-\left[\left(\frac{8}{8}\right)^2+\left(\frac{0}{8}\right)^2\right]$$
$$= 1-[1+0]$$
$$= 0$$

For CarType = Luxury, 1 instance belong to C0, 7 instances belong to C1.

$$\text{Gini Index (CarType = Luxury)} = 1-\left[\left(\frac{1}{8}\right)^2+\left(\frac{7}{8}\right)^2\right]$$
$$= 1-[0.78125]$$
$$= 0.21875$$

$$\text{Weighted Gini Index (Car Type)} = \frac{4}{20}\times 0.375 + \frac{8}{20}\times 0 + \frac{8}{20}\times 0.21875$$
$$= 0.1625$$

(e) Compute the Gini Index for Shirt size attribute using multiway split.

Answer:-

For shirt size, using multiway split, we have 4 possibilities as Small, Medium, Large, and Extralarge

For shirt size = small, 3 instances belong to C0, 2 instances belong to C1.

$$\text{Gini Index (shirt size = small)} = 1-\left[\left(\frac{3}{5}\right)^2+\left(\frac{2}{5}\right)^2\right]$$
$$= 0.48$$

For shirtsize = medium, 3 instances belong to C0, 4 instances belong to C1.

$$\text{Gini Index (shirt size = medium)} = 1-\left[\left(\frac{3}{7}\right)^2+\left(\frac{4}{7}\right)^2\right]$$
$$= 0.48979$$

For shirt size = Large, 2 instances belong to C0, 2 instances belong to C1.

$$\text{Gini Index (shirt size = Large)} = 1-\left[\left(\frac{2}{4}\right)^2+\left(\frac{2}{4}\right)^2\right]$$
$$= 0.5$$

For shirt size = Extra large, 2 instances belong to C0, 2 instances belong to C1.

$$\text{Gini Index (shirt size = Extra large)} = 1-\left[\left(\frac{2}{4}\right)^2+\left(\frac{2}{4}\right)^2\right]$$
$$= 0.5$$

$$\text{Weighted Gini Index (shirt size)} = \frac{5}{20}\times 0.48 + \frac{7}{20}\times 0.48979 + \frac{4}{20}\times 0.5 + \frac{4}{20}\times 0.5$$
$$= 0.49142$$

(f) Which attribute is better Gender, CarType, or shirt size?

Answer:-

Consider the Gini Index as the measure of impurity, we consider the attribute with the lowest Gini index value as the best for splitting criterion.

Gini Index (Gender) = 0.48
Gini Index (CarType) = 0.1625
Gini Index (shirt size) = 0.49142

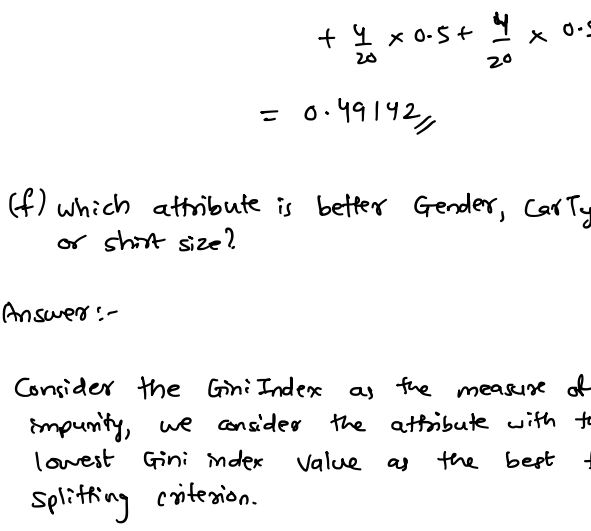
Hence, CarType is the better attribute in comparison to Gender and shirt size.

(g) Explain why customer Id should not be used as the attribute test condition even though it has the lowest Gini.

Answer:-

Even though the Gini for Customer ID is lowest i.e. zero. It cannot be used for attribute test condition because it has a unique value for each instance.

A test condition using customer Id will accurately classify every instance in the training data, it cannot be used on new test instances of test data that the model has never seen / learnt about how to classify.



3. Consider the training examples shown in Table 3.6 for a binary classification problem.

Table 3.6. Data set for Exercise 3.

Instance	a1	a2	a3	Target Class
1	T	T	1.0	+
2	T	T	6.0	+
3	T	F	5.0	-
4	F	F	4.0	+
5	F	T	7.0	-
6	F	T	3.0	-
7	F	F	8.0	-
8	T	F	7.0	+
9	F	T	5.0	-

(a) What is the entropy of this collection of training examples with respect to the class attribute?

Answer:-

Entropy as a measure of impurity is given as

$$\text{Entropy} = -\sum_{i=0}^{c-1} p_i(t)\log_2 p_i(t)$$

For the overall collection, 4 instances belong to class '+' and 5 instances belong to class '-'.

c = 2

$$\text{Entropy} = -\sum_{i=0}^{1} p_i(t)\cdot\log_2 p_i(t)$$
$$= -\left[p_0(t)\cdot\log_2 p_0(t) + p_1(t)\cdot\log_2 p_1(t)\right]$$
$$= -\left[\frac{4}{9}\cdot\log_2\left(\frac{4}{9}\right)+\frac{5}{9}\cdot\log_2\left(\frac{5}{9}\right)\right]$$
$$= -\left[-0.99107\right]$$
$$= 0.99107$$

(b) What are the information gains of a1 and a2 relative to these training examples?

Answer:-

Entropy for a1, when a1 = T, 3 instances belong to '+' & 1 instance belong to '-'

$$\text{Entropy (a1 = T)} = -\left[\frac{3}{4}\log_2\frac{3}{4}+\frac{1}{4}\log_2\frac{1}{4}\right]$$
$$= 0.81127$$

Entropy (a1 = F), 1 instance belong to '+' & 4 instances belong to '-'.

$$\text{Entropy (a1 = F)} = -\left[\frac{1}{5}\log_2\frac{1}{5}+\frac{4}{5}\log_2\frac{4}{5}\right]$$
$$= 0.72192$$

$$\text{Entropy (a1)} = \frac{4}{9}\times 0.81127 + \frac{5}{9}\times 0.72192$$
$$= 0.76163$$

Δ info (Information Gain) for a1

$$= \text{Entropy (root or parent node)} - \text{Entropy (a1)}$$
$$= 0.99107 - 0.76163$$
$$= 0.22944$$

$$\text{Entropy }(a_2) = \frac{5}{9}\times\text{Entropy (a2 = T)} + \frac{4}{9}\times\text{Entropy (a2 = F)}$$
$$= \frac{5}{9}\left[-\frac{2}{5}\log_2\left(\frac{2}{5}\right)-\frac{3}{5}\log_2\left(\frac{3}{5}\right)\right] + \frac{4}{9}\left[-\frac{2}{4}\log_2\left(\frac{2}{4}\right)-\frac{2}{4}\log_2\left(\frac{2}{4}\right)\right]$$
$$= \frac{5}{9}\left[0.97095\right]+\frac{4}{9}\left[1\right]$$
$$= 0.98386$$

Δ info (Information gain) for a2 =

$$\text{Entropy (root or parent node)} - \text{Entropy (a2)}$$
$$= 0.99107 - 0.98386$$
$$= 0.00721$$

(c) For a3, which is a continuous attribute, compute the information gain for every possible split.

Answer:-

class	+	-	+	-, -	+	+, -	-
a3 sorted value	1.0	3.0	4.0	5.0	6.0	7.0	8.0
split position	2.0	3.5	4.5	5.5	6.5	7.5	8.5

	≤	>	≤	>	≤	>	≤	>	≤	>	≤	>	≤	>
class +	1	3	1	3	2	2	2	2	3	1	4	0	4	0
class -	0	5	1	4	1	4	3	2	3	2	4	1	5	0

$$\text{Entropy (value = 1.0)} = \frac{8}{9}\left[-\frac{3}{8}\log_2\frac{3}{8}-\frac{5}{8}\log_2\frac{5}{8}\right] + \frac{1}{9}\left[-\frac{1}{1}\log_2\frac{1}{1}-0\log_2\frac{0}{1}\right]$$
$$= 0.84838$$

$$\Delta\text{info (value = 1.0)} = 0.99107 - 0.84838 = 0.14269$$

$$\text{Entropy (value = 3.0)} = \frac{7}{9}\left[-\frac{3}{7}\cdot\log_2\frac{3}{7}-\frac{4}{7}\cdot\log_2\frac{4}{7}\right] + \frac{2}{9}\left[-\frac{1}{2}\log_2\frac{1}{2}-\frac{1}{2}\log_2\frac{1}{2}\right]$$
$$= 0.98850$$

$$\Delta\text{info (value = 3.0)} = 0.99107 - 0.98850$$
$$= 0.00257$$

$$\text{Entropy (value = 4.0)} = \frac{6}{9}\left[-\frac{2}{6}\log_2\frac{2}{6}-\frac{4}{6}\log_2\frac{4}{6}\right] + \frac{3}{9}\left[-\frac{2}{3}\log_2\frac{2}{3}-\frac{1}{3}\log_2\frac{1}{3}\right]$$
$$= 0.91829$$

$$\Delta\text{info (value = 4.0)} = 0.99107 - 0.91829$$
$$= 0.07278$$

$$\text{Entropy (value = 5.0)} = \frac{4}{9}\left[-\frac{2}{4}\cdot\log_2\frac{2}{4}-\frac{2}{4}\log_2\frac{2}{4}\right] + \frac{5}{9}\left[-\frac{2}{5}\log_2\frac{2}{5}-\frac{3}{5}\log_2\frac{3}{5}\right]$$
$$= 0.98386$$

$$\Delta\text{info (value = 5.0)} = 0.99107 - 0.98386$$
$$= 0.00721$$

$$\text{Entropy (value = 6.0)} = \frac{3}{9}\left[-\frac{1}{3}\log_2\frac{1}{3}-\frac{2}{3}\log_2\frac{2}{3}\right] + \frac{6}{9}\left[-\frac{3}{6}\log_2\frac{3}{6}-\frac{3}{6}\log_2\frac{3}{6}\right]$$
$$= 0.97276$$

$$\Delta\text{info (value = 6.0)} = 0.99701 - 0.97276$$
$$= 0.01831$$

$$\text{Entropy (value = 7.0)} = \frac{1}{9}\left[-\frac{0}{1}\cdot\log_2\frac{0}{1}-\frac{1}{1}\log_2\frac{1}{1}\right] + \frac{8}{9}\left[-\frac{4}{8}\log_2\frac{4}{8}-\frac{4}{8}\log_2\frac{4}{8}\right]$$
$$= 0.88888$$

$$\Delta\text{info (value = 7.0)} = 0.99701 - 0.88888$$
$$= 0.10219$$

$$\text{Entropy (value = 8.0)} = 0 + \frac{9}{9}\left[-\frac{4}{9}\cdot\log_2\frac{4}{9}-\frac{5}{9}\log_2\frac{5}{9}\right]$$
$$= 0.99107$$

$$\Delta\text{info (value = 8.0)} = 0.99107 - 0.99107$$
$$= 0$$

Since the information gain is highest at value = 1.0, It is the best splitting condition for a3.

(d) What is the best split among a1, a2, and a3 according to information gain?

Answer:-

Considering entropy as the measure of impurity, we want to maximize information gain.

Δinfo (a_1) = 0.22944
Δinfo (a_2) = 0.00721
Δinfo (a_3) = 0.14269

As the information gain is highest for a1, a1 is the best split among a1, a2, and a3.

(e) What is the best split among a_1 & a_2 according to misclassification error rate?

Answer:-

Misclassification error rate is given as

$$C.E = 1-\max_i(p_i(t))$$

For a_1 = T,

$$C.E = 1-\max\left(\frac{1}{4},\frac{3}{4}\right) = 1-\frac{3}{4} = \frac{1}{4}$$

For a_1 = F,

$$C.E = 1-\max\left(\frac{1}{5},\frac{4}{5}\right) = 1-\frac{4}{5} = \frac{1}{5}$$

$$C.E(a_1) = \frac{4}{9}\left(\frac{1}{4}\right)+\frac{5}{9}\left(\frac{1}{5}\right)$$
$$= \frac{1}{9}+\frac{1}{9} = \frac{2}{9}$$

For a_2 = T,

$$C.E = 1-\max\left(\frac{2}{5},\frac{3}{5}\right) = 1-\frac{3}{5} = \frac{2}{5}$$

For a_2 = F,

$$C.E = 1-\max\left(\frac{2}{4},\frac{2}{4}\right) = 1-\frac{2}{4} = \frac{2}{4}$$

$$C.E(a2) = \frac{5}{9}\times\frac{2}{5}+\frac{4}{9}\times\frac{2}{4} = \frac{2}{9}+\frac{2}{9}$$
$$= \frac{4}{9}$$

We want lowest misclassification error rate and this is given by a1.
Hence, a1 is a better split among a1 and a2 according to misclassification error rate.

(f) What is the best split among a1 and a2 according to Gini Index?

Answer:-

$$\text{Gini Index (a1 = T)} = 1-\left[\left(\frac{1}{4}\right)^2+\left(\frac{3}{4}\right)^2\right]$$
$$= 0.375$$

$$\text{Gini Index (a1 = F)} = 1-\left[\left(\frac{1}{5}\right)^2+\left(\frac{4}{5}\right)^2\right]$$
$$= 0.32$$

$$\text{Gini Index (a1)} = \frac{4}{9}\times 0.375 + \frac{5}{9}\times 0.32$$
$$= 0.344$$

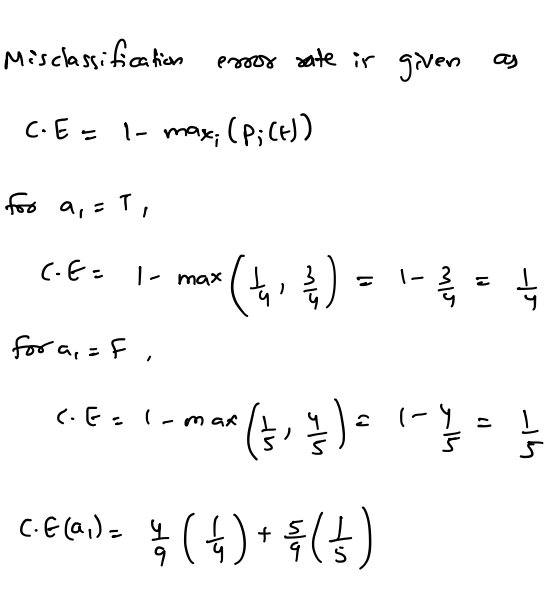
$$\text{Gini Index (a2 = T)} = 1-\left[\left(\frac{2}{5}\right)^2+\left(\frac{3}{5}\right)^2\right]$$
$$= 0.48$$

$$\text{Gini Index (a2 = F)} = 1-\left[\left(\frac{2}{4}\right)^2+\left(\frac{2}{4}\right)^2\right]$$
$$= 0.5$$

$$\text{Gini Index (a2)} = \frac{5}{9}\times 0.48 + \frac{4}{9}\times 0.5$$
$$= 0.488$$

On comparing Gini index (a1) & Gini Index (a2), we want the lowest gini Index. That is provided by a1.

Hence a1 is a better split among a1 and a2 according to Gini Index.



5. Consider the following data set for a binary class problem.

A	B	Class Label
T	F	+
T	T	+
T	T	+
T	F	-
T	T	+
F	F	-
F	F	-
F	F	-
T	T	-
T	F	-

(a) Calculate the information gain when splitting on A and B, which attribute would the decision tree induction algorithm choose?

Answer:-

$$\text{Entropy (A)} = \frac{7}{10}\left[-\frac{4}{7}\log_2\frac{4}{7}-\frac{3}{7}\log_2\frac{3}{7}\right] + \frac{3}{10}\left[-\frac{0}{3}\log_2\frac{0}{3}-\frac{3}{3}\log_2\frac{3}{3}\right]$$
$$= \frac{7}{10}\times(0.98522) + \frac{3}{10}(0)$$
$$= 0.68965$$

$$\text{Entropy (B)} = \frac{4}{10}\left[-\frac{3}{4}\log_2\frac{3}{4}-\frac{1}{4}\log_2\frac{1}{4}\right] + \frac{6}{10}\left[-\frac{1}{6}\log_2\frac{1}{6}-\frac{5}{6}\log_2\frac{5}{6}\right]$$
$$= \frac{4}{10}\times(0.81127) + \frac{6}{10}\times(0.6500)$$
$$= 0.71450$$

$$\text{Entropy (root or parent node)} = -\frac{4}{10}\log_2\frac{4}{10}-\frac{6}{10}\log_2\frac{6}{10}$$
$$= 0.97095$$

$$\Delta\text{info (A)} = \text{Entropy (root)} - \text{Entropy (A)}$$
$$= 0.97095 - 0.68965$$
$$= 0.2813$$

$$\Delta\text{info (B)} = \text{Entropy (root)} - \text{Entropy (B)}$$
$$= 0.97095 - 0.71450$$
$$= 0.25645$$

As the information gain is to be maximized by the decision tree induction algorithm, it will choose attribute A for splitting criterion.

(b) Calculate the gain in Gini Index when splitting on A and B. Which attribute would the decision tree induction choose?

Answer:-

$$\text{Gini Index (A)} = \frac{7}{10}\times\text{Gini Index (A = T)} + \frac{3}{10}\times\text{Gini Index (A = F)}$$
$$= \frac{7}{10}\times\left(1-\left[\left(\frac{4}{7}\right)^2+\left(\frac{3}{7}\right)^2\right]\right) + \frac{3}{10}\left(1-\left[\left(\frac{0}{3}\right)^2+\left(\frac{3}{3}\right)^2\right]\right)$$
$$= \frac{7}{10}(0.48979)+\frac{3}{10}(0)$$
$$= 0.34285$$

$$\text{Gini Index (B)} = \frac{4}{10}\times\text{Gini Index (B = T)} + \frac{6}{10}\times\text{Gini Index (B = F)}$$
$$= \frac{4}{10}\left(1-\left[\left(\frac{3}{4}\right)^2+\left(\frac{1}{4}\right)^2\right]\right) + \frac{6}{10}\left(1-\left[\left(\frac{1}{6}\right)^2+\left(\frac{5}{6}\right)^2\right]\right)$$
$$= \frac{4}{10}\times 0.375 + \frac{6}{10}\times 0.2777$$
$$= 0.31667$$

$$\text{Gini Index (root)} = 1-\left[\left(\frac{4}{10}\right)^2+\left(\frac{6}{10}\right)^2\right]$$
$$= 0.48$$

$$\Delta\text{gain (A)} = \text{Gini Index (root)} - \text{Gini Index (A)}$$
$$= 0.48 - 0.34285$$
$$= 0.13715$$

$$\Delta\text{gain (B)} = \text{Gini Index (root)} - \text{Gini Index (B)}$$
$$= 0.48 - 0.31667$$
$$= 0.16333$$

Since, larger the Δ gain, better is the split. Hence, B will be choosen by the decision tree induction algorithm for split.

(c) Figure 3.11 shows that entropy and Gini Index are both monotonially increasing on the range [0, 0.5] and they are both monotonically decreasing on the range [0.5, 1]. Is it possible that information gain and the gain in index favor different attributes? Explain.

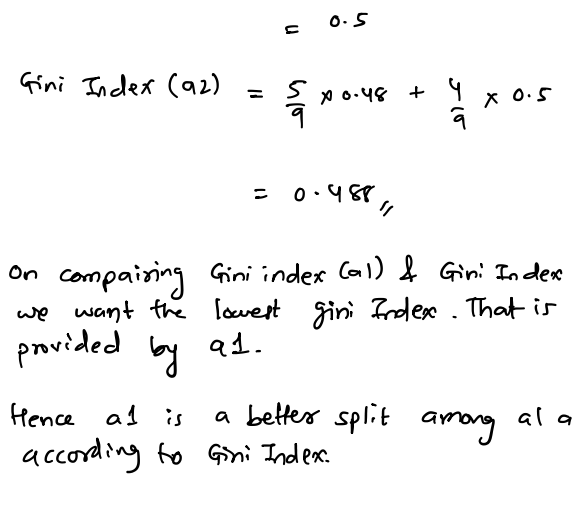


Figure 3.11.
Comparison among the impurity measures for binary classification problems.

Answer:-

Comparing the results of Q.5(a) and Q.5(b), we can witness that based on information gain attribute A will be chosen for the splitting criterion.
While, based on gain in index, attribute B will be chosen for the splitting criterion.

Even though the entropy and gini index increase and decrease monotonically in their respective ranges, the gains they provide showcase a different trend as per the results obtained in Q5(a) & Q5(b).

Exercise 1.2 - Tan (chapter 4)

18. Consider the task of building a classifier from random data, where the attribute labels are generated randomly irrespective of the class labels. Assume the data set contains instances from two classes "+" & "-". Half of the data set is used for training & half of the data set is used for testing.

(a) Suppose there are an equal numbers of positive and negative instances in the data, and the decision tree classifier predicts every test instance to be positive. What is the expected error rate of the classifier on the test data?

Answer:-

Let's consider the numbers of instances to be as N.
$\frac{N}{2}$ instances are used for training and $\frac{N}{2}$ instances for testing.

$\frac{N}{2}$ instances class labels are "+" & $\frac{N}{2}$ instances class labels are "-".

The classifier predicts every instance of test data to be positive.
misclassification error rate

$$= 1-\max\left(\frac{1}{2},\frac{1}{2}\right)$$
$$= \frac{1}{2}$$

Hence, the classifier will misclassify 50% of the data instances.

(b) Repeat the previous analysis assuming that the classifier predicts each test instance to be positive with probability of 0.8 and negative class with probability of 0.2.

Answer:-

$\frac{N}{2}$ instances are "+" & $\frac{N}{2}$ instances are "-"

Classifier predicts "+" with a probability of 0.8 i.e. 20% of "+" are misclassified

Similarly, 80% of "-" are misclassified.

Total misclassification error rate

$$= \left(20\% \text{ of } \frac{N}{2} + 80\% \text{ of } \frac{N}{2}\right)/N$$
$$= \frac{\frac{20}{100}\times\frac{N}{2}+\frac{80}{100}\times\frac{N}{2}}{N} = \frac{\frac{100}{100}\times\frac{N}{2}}{N} = \frac{1}{2}$$
$$= 0.5$$

Hence, 50% of the data instances are misclassified.

(c) Suppose two thirds of data belong to positive class and remaining one third belong to negative class. What is the expected error of a classifier that predicts every test instance to be positive?

Answer:-

$\frac{2}{3}N$ belongs to "+" class & $\frac{1}{3}N$ belongs to "-" class.

$$\text{misclassification error rate} = 1-\max\left(\frac{2}{3},\frac{1}{3}\right)$$
$$= 1-\frac{2}{3} = \frac{1}{3} = 0.33$$

Hence 33% of data instances will be misclassified by the classifier.

(d) Repeat the previous analysis assuming that the classifier predicts each test instance to be positive class with probability of 2/3 & negative class with probability of 1/3.

Answer:-

$\frac{2}{3}N$ belongs to "+" class, $\frac{1}{3}N$ belongs to "-" class

Classifier predicts "+" class with a probability of $\frac{2}{3}$. i.e $\frac{1}{3}$ instances are misclassified

Similarly, $\frac{2}{3}$ instances of "-" are misclassified.

Total misclassification error rate =

$$\frac{\left(\frac{1}{3}\times\frac{2}{3}N+\frac{2}{3}\times\frac{1}{3}N\right)}{N}$$
$$= \frac{\frac{2}{9}N+\frac{2}{9}N}{N}$$
$$= \frac{4N}{9N} = \frac{4}{9} = 0.444$$

Hence, 44.4% of instances will be misclassified.

Exercise 1.3

Multiclass classification.

Using the confusion matrix from multiclass.Rmd notebook (from Lecture 7), create a binary class confusion matrix using the "one-vs-many" strategy for each class. Then, for each class, compute sensitivity, specificity, and precision to two decimal places. Show all work, including the binary class confusion matrices.

Answer:-

Actual → ↓ Predicted	Setosa	Versicolor	Virginica
Setosa	10	0	0
Versicolor	0	10	1
Virginica	0	0	9

For setosa, using one-vs-many approach:-

Actual → ↓ Predicted	Setosa	Versicolor, Virginica
Setosa	10 tp	0 fp
Versicolor, Virginica	0 fn	20 tn

$$\text{sensivity} = \frac{tp}{tp+fn} = \frac{10}{10+0} = 1.00$$

$$\text{specificity} = \frac{tn}{fp+tn} = \frac{20}{0+20} = 1.00$$

$$\text{precision} = \frac{tp}{tp+fp} = \frac{10}{10+0} = 1.00$$

For Versicolor, using one-vs-many approach:-

Actual → ↓ Predicted	Versicolor	Setosa, Virginica
Versicolor	10 tp	1 fp
Setosa, Virginica	0 fn	19 tn

$$\text{sensitivity} = \frac{tp}{tp+fn} = \frac{10}{10+0} = 1.00$$

$$\text{Specificity} = \frac{tn}{fp+tn} = \frac{19}{1+19} = 0.95$$

$$\text{precision} = \frac{tp}{tp+fp} = \frac{10}{10+1} = 0.91$$

For Virginica, using one-vs-many approach:-

Actual → ↓ Predicted	Virginica	Setosa, Versicolor
Virginica	9 tp	0 fp
Setosa, Versicolor	1 fn	20 tn

$$\text{sensitivity} = \frac{tp}{tp+fn} = \frac{9}{9+1} = 0.90$$

$$\text{specificity} = \frac{tn}{tn+fp} = \frac{20}{20+0} = 1.00$$

$$\text{precision} = \frac{tp}{tp+fp} = \frac{9}{9+0} = 1.00$$