PROBLEM 1 (15pts)

1. Classification Rules (5pts)

For the following **formulas** write and use the proper definitions to prove whether they are or they are not **discriminant** or **characteristic rules** in the following dataset DB

CLASSIFICATION DB

O	a1	a2	a3	a4	C
o1	1	1	1	0	1
o2	2	1	2	0	2
o3	0	0	0	0	0
o4	0	0	2	1	0
o5	2	1	1	0	1

Characteristic Rule Definition

Discriminant Rule Definition

Formulas

f1 $a1 = 1 \cap a2 = 1 \Rightarrow C = 1$

f1 is / is not Discriminant Rule because

f2 $C = 1 \Rightarrow a1 = 0 \cap a2 = 1 \cap a3 = 1$

f2 is / is not Characteristic Rule because

f3 $a1 = 1\,. \Rightarrow C = 1$

f3 is / is not Discriminant Rule because

f4 $C = 1 \Rightarrow a1 = 1$

f4 is / is not Characteristic Rule because

f5 $a1 = 2 \cap a2 = 1 \cap a3 = 1 \Rightarrow C = 0$

f5 is / is not Characteristic Rule because

2. (5pts)

Prove that in any classification DB the inverse implication to the discriminant rule is a characteristic rule

3. (5pts)

Given classification DB

Find a simple condition (example) under which the inverse implication to a characteristic rule

is ALWAYS a discriminant rule

PART-2

PROBLEM-1:

1. Given :-

O	a_1	a_2	a_3	a_4	C
o_1	1	1	1	0	1
o_2	2	1	2	0	2
o_3	0	0	0	0	0
o_4	0	0	2	1	0
o_5	2	1	1	0	1

CHARACTERISTIC RULE Definition :- A characteristic formula CLASS ⇒ DESCRIPTION is called a "Characteristic Rule" of the classification dataset DB if it is "true" in DB, i.e when the following holds:-

$\boxed{\{o: \text{DESCRIPTION}\} \cap \{o: \text{CLASS}\} \text{ not } = \text{empty set}}$

where {o: DESCRIPTION} is a set of all records of DB corresponding to the description "DESCRIPTION" & {o: CLASS} is the set of all records of DB corresponding to the description "CLASS"

DISCRIMINANT-RULE Definition :- A discriminant formula DESCRIPTION ⇒ CLASS" is called a "Discriminant Rule" of DB, if it is "True" on DB, i.e the following two conditions hold :-

(i) $\{o : \text{DESCRIPTION}\}$ not = empty set
(ii) $\{o : \text{DESCRIPTION}\}$ included in $\{o : \text{CLASS}\}$

(i) $a_1 = 1 \cap a_2 = 1 \Rightarrow C = 1$

DESCRIPTION $\Rightarrow$ CLASS

$a_1 = 1$

Record	a_2	a_3	a_4	c
O_1	1	1	0	1

$a_1 = 1 \cap a_2 = 1$

Record	a_3	a_4	c
O_1	1	0	1

$C = 1$

Record	a_1	a_2	a_3	a_4
O_1	1	1	1	0
O_5	2	1	1	0

$\{O : a_1 = 1 \wedge a_2 = 1\} = \{O_1\}$

$\Rightarrow \{O : \text{DESCRIPTION}\}$ not empty set

$\{O : C = 1\} = \{O_1, O_5\}$

$\{O : a_1 = 1 \cap a_2 = 1\} = \{O_1\} \subseteq \{O : C = 1\} = \{O_1, O_5\}$

$\Rightarrow \{O : \text{DESCRIPTION}\} \subseteq \{O : \text{CLASS}\}$

$\because$ Since $\{O_1\}$ is a subset of $\{O_1, O_5\}$ $\Rightarrow$ So this is "DISCRIMINANT" rule

Answer

Scanned with CamScanner

(ii) $C=1 \Rightarrow a_1=0 \cap a_2=1 \cap a_3=1$

CLASS $\Rightarrow$ DESCRIPTION

$C=1$

Record	a_1	a_2	a_3	a_4
o_1	1	1	1	0
o_5	2	1	1	0

$a_1=0$

Record	a_2	a_3	a_4	C
o_3	0	0	0	0
o_4	0	2	1	0

$a_1=0 \cap a_2=1$

None

$a_1=0 \cap a_2=1 \cap a_3=1$

None

$\{o: a_1=0 \cap a_2=1 \cap a_3=1\} = \phi$

$\{o: C=1\} = \{o_1, o_5\}$

$\{o: a_1=0 \cap a_2=1 \cap a_3=1\} = \phi \cap \{o: C=1\} = \{o_1, o_5\} = \phi$

$\{o: \text{DESCRIPTION}\} \cap \{o: \text{CLASS}\} = \phi$

$\therefore$ since $\{o: a_1=0 \ \& \ a_2=1 \ \& \ a_3=1\}$ is a empty set, so this "not a CHARACTERISTIC" rule.

Answer

Scanned with CamScanner

(iii) $a_1 = 1 \Rightarrow C = 1$

DESCRIPTION $\Rightarrow$ CLASS

$a_1 = 1$

Record	a_2	a_3	a_4	C
O_1	1	1	0	1

$C = 1$

Record	a_1	a_2	a_3	a_4
O_1	1	1	1	0
O_5	2	1	1	0

$\{O : a_1 = 1\} = \{O_1\}$

$\Rightarrow \{O : \text{DESCRIPTION}\}$ not empty set

$\{O : C = 1\} = \{O_1, O_5\}$

$\{O : a_1 = 1\} = \{O_1\} \subseteq \{O : C = 1\} = \{O_1, O_5\}$

$\Rightarrow \{O : \text{DESCRIPTION}\} \subseteq \{O : \text{CLASS}\}$

$\therefore$ Since $\{O_1\}$ is a subset of $\{O_1, O_5\}$ $\Rightarrow$ This is a "DISCRIMINANT RULE"

Answer.

(iv) $C=1 \Rightarrow a_1 = 1$

CLASS $\Rightarrow$ DESCRIPTION

$C=1$

Record	a_1	a_2	a_3	a_4
O_1	1	1	1	0
O_5	2	1	1	0

$a_1 = 1$

Record	a_2	a_3	a_4	c
O_1	1	1	0	1

$\{o : a_1 = 1\} = \{O_1\}$

$\{o : C = 1\} = \{O_1, O_5\}$

$\{o : a_1 = 1\} = \{O_1\} \cap \{o : C=1\} = \{O_1, O_5\} = \{O_1\}$

$\Rightarrow \{o : \text{DESCRIPTION}\} \cap \{o : \text{CLASS}\}$ not = emptyset

$\therefore$ Since $\{O_1\} \cap \{O_1, O_5\} = \{O_1\}$ which is not empty set

$\therefore$ This is a CHARACTERISTIC Rule

Answer.

(v) $a_1 = 2 \cap a_2 = 1 \cap a_3 = 1 \Rightarrow C = 0$

DESCRIPTION $\Rightarrow$ CLASS

$a_1 = 2$

Record	a_2	a_3	a_4	C
O_2	1	2	0	2
O_5	1	1	0	1

$a_1 = 2 \cap a_2 = 1 \Rightarrow \{O_2, O_5\}$.

$a_1 = 2 \cap a_2 = 1 \cap a_3 = 1 \Rightarrow \{O_5\}$

$C = 0$

Record	a_1	a_2	a_3	a_4
O_3	0	0	0	0
O_4	0	0	2	1

$\{O : a_1 = 2 \cap a_2 = 1 \cap a_3 = 1\} = \{O_5\}$

$\Rightarrow \{O : \text{DESCRIPTION}\}$ not empty set

$\{O : C = 0\} = \{O_3, O_4\}$

$\Rightarrow \{O : a_1 = 2 \cap a_2 = 1 \cap a_3 = 1\} = \{O_5\}$ is not a subset of

$\{O : C = 0\} = \{O_3, O_4\}$.

$\therefore$ Since $\{O_5\}$ is not subset of $\{O_3, O_4\}$, This is

"not a DISCRIMINANT" rule

Answer.

And since the format is wrong i.e DESCRIPTION ⇒ CLASS,

It is "not a characteristic Rule". For the Characteristic

Rule, the format should be CLASS ⇒ DESCRIPTION.

∴ This is neither CHARACTERISTIC RULE (due to wrong format)

nor DISCRIMINANT RULE (due to conditions not satisfying)

Answer.

2. Prove that in any classification DB, the inverse implication to the discriminant rule is a characteristic rule ??

Answer:-

By definition, for any data base DB:

DESCRIPTION $\Rightarrow$ CLASS

is a discriminant rule iff the below conditions holds true:

1) {o: DESCRIPTION} is not empty
2) {o: DESCRIPTION} is included in {o: CLASS}

We know that for any non-empty sets A, B, if A is included in B, then their intersection is non-empty.

Hence, {o: DESCRIPTION} intersection with {o: CLASS} is not empty and by definition, inverse application i.e,

CLASS $\Rightarrow$ DESCRIPTION

is a CHARACTERISTIC RULE

Example RULE: $a_1 = 1 \Rightarrow c = 1$

DESCRIPTION $\Rightarrow$ CLASS

DB

O	a_1	a_2	c
o_1	1	1	1
o_2	2	1	2
o_3	2	1	1

Since $\{o: a_1 = 1\} = \{o_1\}$ is a subset of $\{o: c = 1\} = \{o_1, o_3\}$, This is a Discriminant Rule

INVERSE IMPLICATION:- $c = 1 \Rightarrow a_1 = 1$

[CLASS $\Rightarrow$ DESCRIPTION]

~~Since {o1,o3}~~ Since $\{o_1\} \cap \{o_1, o_3\} = \{o_1\}$ which is not empty set, It is a Characteristic Rule

Hence proved

3. Given a classification DB, find a simple condition, under which inverse implication to a characteristic rule is always a discriminant rule

Answer: By definition, for any database DB:

CLASS ⇒ DESCRIPTION

is a characteristic rule, if below condition holds good:

1) {o: DESCRIPTION} ∩ {o: CLASS} not = empty set

We know that for any sets A, B, if A ∩ B is not an empty set, then neither of A and B are empty sets —— condition (1)

⇒ A and B are non-empty sets.

For inverse implication rule ⇒ $\boxed{\text{DESCRIPTION} \Rightarrow \text{CLASS}}$ to be a DISCRIMINANT RULE, as per the definition, we have two conditions to be satisfied, i.e

Condition 1: {o: DESCRIPTION} is not empty

Condition 2: {o: DESCRIPTION} is included in {o: CLASS}

From (1), condition 1 is already met. Therefore, condition 2 needs to be satisfied for this to be true.

⇒ Inverse implication to a characteristic rule is always a Discriminant Rule if {o: DESCRIPTION} is included in {o: CLASS}.

Scanned with CamScanner

Example:-

DB

O	a_1	a_2	c
O_1	1	1	1
O_2	2	1	2
O_3	2	1	1

Rule: $C=1 \Rightarrow a=1$ is a Characteristic Rule since.

$\{o: c=1\} = \{o_1, o_3\} \cap \{o: a=1\} = \{o_1\} = \{o_1\}$ is not empty

Inverse Implication: $a=1 \Rightarrow C=1$ also holds good here, i.e

It is a Discriminant rule because. $\{o: a=1\} = \{o_1\}$ is included in $\{o: c=1\} = \{o_1, o_3\}$

$\therefore$ $\{o:$ DESCRIPTION$\} \subseteq \{o:$ CLASS$\}$ For inverse implication of Characteristic to be always a discriminant rule.