cse352 ARTIFICIAL INTELLIGENCE MIDTERM SOLUTIONS Spring 2022
(65pts) +(5 extra pts)

PART 1: SHORT QUESTIONS (10pts)

Q1 (5pts)

1. (2pts) Define classification data.

Solution

Classification data is a data table (with key attribute removed) with one attribute distinguished as a class attribute. Class attribute must have a small number discrete values. We distinguish attribute part of the record and call it sample/ tuple and class part as determined by the class label attribute. Each tuple/sample is assumed to belong to a predefined class, as determined by the class label attribute.

2. (3pts) Define shortly a **Classifier**

Solution

CLASSIFIER is a ”black box” that had been trained and tested and evaluated. It serves to classify unknown records; i.e tuples for which the class attribute is unknown

Q2 (5pts) The **termination conditions** for Decision Tree Algorithm are as follows.

Solution

T1. All records (samples) for the given **node N** belong to the same class. We convert node into a **leaf** and label it with this class.

T2. There are no remaining attributes on which the samples may be further partitioned - a **leaf** is created with **Majority Vote** for training data D.

T3. There is no records (samples) left a **leaf** is created with **Majority Vote** for training data D.

4. Majority Vote involves converting **node N** into a **leaf** and labeling it with the most common class in the TRAINING data D, which is a set of training tuples and their associated class labels.

PART 2: PROBLEMS · (60 pts)

PROBLEM 1 (10pts) Translation to Predicate Logic

Translate to Predicate Logic the following statement

" Every student likes good grades"

Solution

Observe that "good grades" is a short for "all good grades" and the statement has a hidden universal quantifier "for all"; i.e. we re-write, and translate the statement stated in a more logical way as

" Every student likes all good grades"

Domain is a set $X \neq \emptyset$

s1 Predicates are:

One argument predicates: $S(x)$ - for x is a student, $D(y)$ - for y is good, and $G(y)$ - for y is a grade

Two argument predicate: $L(x, y)$ - for x likes y

s2 Restricted Quantifiers are: only universal $\forall_{S(x)}$ - for every student, $\forall_{(D(y) \cap G(y))}$ - for all good grades

s3 Restricted Quantifiers Formula is: $\forall_{S(x)} \forall_{(D(y) \cap G(y))} L(x, y)$

s4 Logic Formula is:

$$\forall x\, (S(x) \Rightarrow \forall y\, ((D(y) \cap G(y)) \Rightarrow L(x, y)))$$

s5 Logic Formula written in **AI intended interpretation** is:

$$\forall x\, (student(x) \Rightarrow \forall y\, ((good(y) \cap grade(y)) \Rightarrow likes(x, y)))$$

PROBLEM 2 (10pts) Resolution

1. (5pts) Find **all resolvents** of the set **CL** $= \{ C1, C2 \}$ of clauses for $C1 = \{ a, b, c, \neg d \}$ and $C2 = \{ \neg a, \neg b, d \}$.

It means locate all Complementary Pairs in **CL** and resolve them.

Solution

$C1(a)$ and $C2(\neg a)$ resolves on $\{b, c, \neg d, \neg b, d\}$

$C1(b)$ and $C2(\neg b)$ resolves on $\{a, c, \neg d, \neg a, d\}$

$C1(\neg d)$ and $C2(d)$ resolves on $\{a, b, c, \neg a, \neg b\}$

2. (5pts) Use the Resolution Completeness to prove the set **CL** $= \{ C1, C2, C3 \}$ is **satisfiable**, where $C1 = \{ a, b \}$, $C2 = \{ \neg a, c \}$ and $C3 = \{ \neg b, c \}$.

Solution

There are only TWO possible derivations.

D1

$\{a, b\}$, $\{\neg a, c\}$, $\{\neg b, c\}$

$\{b, c\}$, $\{\neg b, c\}$

$\{c\}$

D2

$\{a, b\}$, $\{\neg a, c\}$, $\{\neg b, c\}$

$\{a, c\}$, $\{\neg a, c\}$

$\{c\}$

This proves that we can never get a derivation of $\{\}$.

Hence by the **Resolution Completeness Theorem** we proved that the set **CL** = $\{C1, C2, C3\}$ of clauses is **satisfiable**.

PROBLEM 3 (10pts) Rule Based Systems

Here is a small set of RULES proposed for a simple rule-based system **S** for dealing with a bank.

R1 IF savings are not adequate, THEN invest in savings

R2 IF savings are adequate AND income is adequate THEN invest in stocks

R3 IF there is no children THEN savings are adequate

1. (4pts) Conceptualize the rules **R1 - R3** in **propositional convention** that uses negation.

Solution

My propositional variables are:

A - for: " savings are adequate "

B - for: " invest in savings"

C - for: "income is adequate "

D - for: " invest in stocks "

E - for: " there are children "

My rules are:

R1 $\neg A \Rightarrow B$

R2 $A \cap C \Rightarrow D$

R3 $\neg E \Rightarrow A$

2. (6pts) **Follow** the steps below to conceptualize the rules

R1 IF savings are not adequate, THEN invest in savings

R2 IF savings are adequate AND income is adequate THEN invest in stocks

R3 IF there is no children THEN savings are adequate

in **predicate convention**: *attribute(x, attribute value)*.

s1 (2pt) DEFINE all needed ATTRIBUTES and their values.

USE the intended interpretation NAMES for the ATTRIBUTES

Solution

I use the intended interpretation names for ATTRIBUTES - you can use your own names

The ATTRIBUTES and their VALUES are:

savings with values adequate , not adequate

investSavings with values: yes, no

investStocks with values: yes, no

Income with values: adequate, not adequate

hasChildren with values: yes, no

s2 (2pt) WRITE the RULES

Solution

R1 $saving(x, notadequate) \Rightarrow investSaving(x, yes)$

R2 $saving(x, adequate) \cap income(x, adequate) \Rightarrow investStocks(x, yes)$

R3 $hasChildren(x, no) \Rightarrow saving(x, adequate)$

s3 (2pt) WRITE a database TABLE with your own example of any **2 records** describing some facts in **S**

Solution

There are 4 of my records, yours can be different!

Obj	savings	investSavings	Income	investStocks	hasChildren
o_1	adequate	yes	not adequate	no	yes
o_2	not adequate	no	adequate	yes	no
o_3	adequate	yes	adequate	no	yes
o_4	not adequate	yes	not adequate	no	no

PROBLEM 4 (10pts) Classification Rules

For the following **formulas** write and use the proper definitions te 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

to prove whether they are or they are not **discriminant** or **characteristic rules** in the dataset DB.

Formulas (2pt) each

Solution

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

f1 is a Discriminant Rule because it is a Discriminant Formula and $\{o_1\}$ is a subset of $\{o_1, o_5\}$

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

f2 is a Characteristic Formula but **is not** a Characteristic Rule because $\{o : a1 = 0 \cap a2 = 1 \cap a3 = 1\} = \emptyset$ and $\{o_1\} \neq \emptyset$

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

f3 is a Discriminant Rule because it is a Discriminant Formula and $\{o_1\}$ is a subset of $\{o_1, o_5\}$

f4 $C = 1 \Rightarrow a1 = 1$

f4is a Characteristic Rule because it is a Characteristic Formula and $\{o_1, o_5\} \cap \{o_1\} \neq \emptyset$

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

f5 is **not** Characteristic Rule because it is **not** a Characteristic **Formula**

Remark We used the following definitions

D1. A characteristic **formula** $CLASS \Rightarrow DESCRIPTION$ is called a **Characteristic Rule** in the classification dataset **D** if and only if it is TRUE in **D** i.e. when the following holds

$$\{o \in \mathbf{D} : DESCRIPTION\} \cap \{o \in \mathbf{D} : CLASS\} \neq \emptyset$$

D2. A discriminant **formula** $DESCRIPTION \Rightarrow CLASS$ is called a **Discriminant Rule** in the classification dataset **D** if and only if it is TRUE in **D** i.e. when the following conditions hold

1. $\{o \in \mathbf{D}\ DESCRIPTION\} \neq \emptyset$

2. $\{o \in \mathbf{D}\ DESCRIPTION\} \subseteq \{o \in \mathbf{D}\ CLASS\}$

PROBLEM 5 (20pts) Classification by Decision Tree

Given a Classification DB

O	a1	a2	C
o1	1	1	1
o2	0	0	0
o3	0	1	0
o4	0	0	0
o5	1	1	1
o6	1	1	0
o7	0	0	0
o8	1	0	1

GOAL: use the above DB and **two - fold cross validation holdout** to build a CLASSIFIER using the Decision Tree BASIC Algorithm with **a1** as the root.

Remember applying correctly the TERMINATION CONDITIONS.

Remember that division into 2-folds is an ARBITRARY partition of records into 2 disjoint sets; so there may be many answers depending on the partitions.

Build your CLASSIFIER using the following fold **F**

F = $\{o1, o2, o3, o4\}$ for training - and rest for testing.

Here are steps you must follow.

STEP 1 (5pts) Build (draw) the Decision Tree with **a1** as the root .

The correct Tree in unique. No partial credit

Solution

Observe that the MAJORITY VOTING class is $C = 0$ - as it is a majority class in the Training set.

FOLD **F** = $\{o1, o2, o3, o4\}$ is used for **TRAINING**

a1	a2	C
1	1	1
0	0	0
0	1	0
0	0	0

TRAIN Tree with **root** $a1$ is:

$a1 = 1$ **branch** is:

a2	C
1	1

By Termination Condition **T1.**, the **leaf** on this branch is $C = 1$ and we have a rule $R1 : a1(x, 1) \Rightarrow C(x, 1)$

$a1 = 0$ **branch** is:

a2	C
1	0
1	0

By **T1**, the **leaf** on this branch is $C = 0$ and we have $R2 : a1(x, 0) \Rightarrow C(x, 0)$

STEP 2 (5pts) Write the set of RULES resulting from your tree in a predicate form.

Solution

The set of rules is:

$$RULES = \{ \ R1 : a1(x, 1) \Rightarrow C(x, 1), \quad R2 : a1(x, 0) \Rightarrow C(x, 0) \ \}$$

STEP 3 (5pts) Evaluate your **rules accuracy** and **predictive accuracy**

Solution

FOLD **F1** = $\{o5, o6, o7, o8\}$ is used for **TESTING**

a1	a2	C
1	1	1
1	1	0
0	0	0
1	0	1

Tuple o5 is classified by R1, o6 is misclassified by R1, o7 is classified by R2, o8 is classified by R1.

Predictive accuracy for **f1** = 75% .

Rules Accuracy - testing with Training Data $\{o1, o2, o3, o4\}$

Tuple o1 is classified by R1, o2 is classified by R1, o3 is classified by R1, o4 is classified by R2

Rules accuracy for **F1** is 100%

STEP 4 (5pts) Write your answer as " **My Classifier is:** " and Justify your decision.

Solution 1

Predictive accuracy for obtained rules is 75% and Rules accuracy is 100% and we **ACCEPT** the results as a classifier, hence **"My Classifier "** is

$$\mathbf{C} = \{\ R1:\ a1(x,1) \Rightarrow C(x,1), \quad R2:\ a1(x,0) \Rightarrow C(x,0)\ \}$$

Solution 2

Even of the rules accuracy is 100%, I fund predictive accuracy too low and decided to TRAIN with another training sets with hope to would obtain a higher predictive accuracy, so my answer at this stage is **"there is no classifier"**.