Week 4: Decision Rules and Classification Rules

CS286: Topics in Intelligent Systems

Katarzyna Tarnowska
Ph.D.

The Department of Computer Science
College of Science
San Jose State University

9/9/2020

Agenda

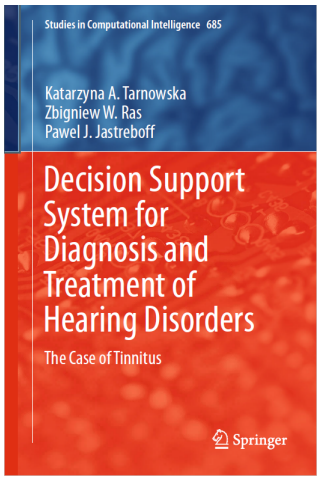
1. Introduction
   - Readings
   - Definition
2. Case study - medical decision making
   - Tinnitus treatment
   - Clinical decision rules
3. Algorithms
   - Rule discovery
   - Activity

Agenda

Agenda

Required readings

1. **Chapter 4.2.-4.3.: Decision Rules. Classification Rules.** in Decision Support System for Diagnosis and Treatment of Hearing Disorders. The Case of Tinnitus. OR
2. **Chapter 4.1.2: Classification** in Recommender System for Improving Customer Loyalty.



Agenda

Decision rule - definition and characteristics

Decision rule

A rule in the form $(\phi \rightarrow \delta)$, where ϕ is called *antecedent* (or assumption) and δ is called *descendant* (or thesis) of the rule.

Atomic decision rule

Rule in form: $r \equiv (a_{i_1} = v_1) \wedge \ldots \wedge (a_{i_m} = v_m) \implies (dec = k)$

Characteristics

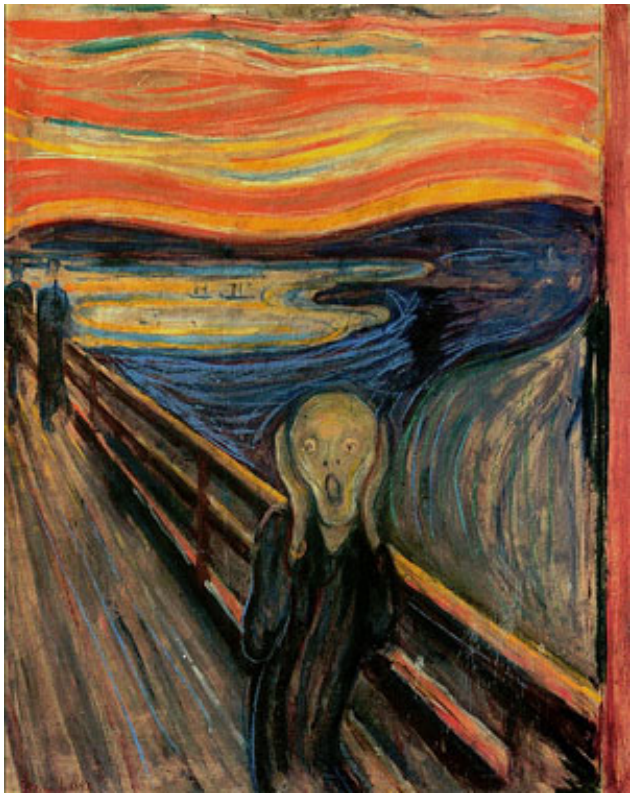
Support(r) - number of objects matching the rule's antecedent
Confidence(r) - relative number of objects matching both rule's antecedent and descendant.

Agenda

Agenda

Tinnitus

- "Ringing in the ears", "phantom auditory perception"
- Affects a significant portion of the population nowadays
- Deteriorates quality of life
- Causes often not clear
- No cure, treatment methods ineffective



Tinnitus retraining therapy (TRT) protocol

Sound Therapy

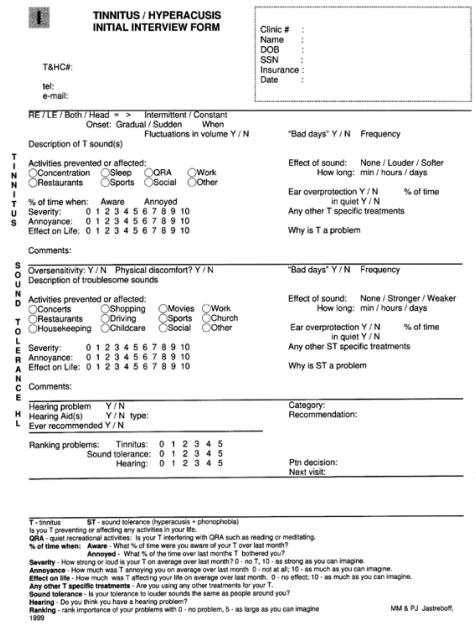
- Sound stimulation provided to the auditory system aims at decreasing tinnitus signal
- Sound generator devices must be fitted

Counseling

- Aims at reclassifying tinnitus to neutral stimuli ("retrain" conditioned reflexes)
- Delivered as individual therapy

Data collection

Structured interview (questionnaires) + audiological assessment



TINNITUS / HYPERACUSIS
INITIAL INTERVIEW FORM

Clinic #
Name
DOB
SSN
Insurance
Date

T&HC#:
tel:
e-mail:

RE / LE / Both / Head = > Intermittent / Constant
Onset: Gradual / Sudden When
Fluctuations in volume Y / N

Description of T sound(s)

"Bad days" Y / N Frequency

Activities prevented or affected:
Concentration Sleep QRA Work
Restaurants Sports Social Other

Effect of sound: None / Louder / Softer
How long: min / hours / days

% of time when: Aware Annoyed
in quiet Y / N
Severity: 0 1 2 3 4 5 6 7 8 9 10
Annoyance: 0 1 2 3 4 5 6 7 8 9 10
Effect on Life: 0 1 2 3 4 5 6 7 8 9 10

Ear overprotection Y / N % of time
Any other T specific treatments

Why is T a problem

Comments:

SOUND TOLERANCE HL

Oversensitivity: Y / N Physical discomfort? Y / N
Description of troublesome sounds

"Bad days" Y / N Frequency

Activities prevented or affected:
Concerts Shopping Movies Work
Restaurants Driving Sports Church
Housekeeping Childcare Social Other

Effect of sound: None / Stronger / Weaker
How long: min / hours / days

Ear overprotection Y / N % of time
in quiet Y / N
Severity: 0 1 2 3 4 5 6 7 8 9 10
Annoyance: 0 1 2 3 4 5 6 7 8 9 10
Effect on Life: 0 1 2 3 4 5 6 7 8 9 10

Any other ST specific treatments

Why is ST a problem

Comments:

Hearing problem Y / N
Hearing Aid(s) Y / N type:
Ever recommended Y / N

Category:
Recommendation:

Ranking problems: Tinnitus: 0 1 2 3 4 5
Sound tolerance: 0 1 2 3 4 5
Hearing: 0 1 2 3 4 5

Ptn decision:
Next visit:

T - tinnitus ST - sound tolerance (hyperacusis + phonophobia)
Is you T preventing or affecting any activities in your life.
QRA - quiet recreational activities: Is your T interfering with QRA such as reading or meditating.
% of time when: Aware - What % of time were you aware of your T over last month?
Annoyed - What % of the time over last months T bothered you?
Severity - How strong or loud is your T on average over last month? 0 - no T, 10 - as strong as you can imagine.
Annoyance - How much was T annoying you on average over last month 0 - not at all; 10 - as much as you can imagine.
Effect on life - How much was T affecting your life on average over last month. 0 - no effect; 10 - as much as you can imagine.
Any other T specific treatments - Are you using any other treatments for your T.
Sound tolerance - Is your tolerance to louder sounds the same as people around you?
Hearing - Do you think you have a hearing problem?
Ranking - rank importance of your problems with 0 - no problem, 5 - as large as you can imagine
MM & PJ Jastreboff, 1999

FU
TINNITUS / HYPERACUSIS
FOLLOW-UP INTERVIEW FORM

CATEGORY:
Date of Init. couns.
Date of instr. fit.
SG:
HA:
FUQ #:
Month #:

Clinic #
Name
DOB
SSN
Insurance
Date

T&HC#:
tel:

"Bad days" Y / N Frequency
Are they: as frequent Y /N as bad Y / N

Activities prevented or affected: Changes: Y / N
Concentration Sleep QRA Work

Effect of sound: None / Louder / Softer

Is you T preventing or affecting any activities in your life.
QRA - quiet recreational activities: Is your T interfering with QRA such as reading or meditating. ◐ - an activity affected as first visit ◑ - an activity affected as for today
% of time when: Aware - What % of time were you aware of your T over last month?
Annoyed - What % of the time over last months T bothered you?
Severity - How strong or loud is your T on average over last month? 0 - no T, 10 - as strong as you can imagine.
Annoyance - How much was T annoying you on average over last month 0 - not at all; 10 - as much as you can imagine.
Effect on life - How much was T affecting your life on average over last month. 0 - no effect; 10 - as much as you can imagine.
Any other T specific treatments - Are you using any other treatments for your T.
Sound tolerance - Is your tolerance to louder sounds the same as people around you?
Hearing - Do you think you have a hearing problem?
Ranking - rank importance of your problems with 0 - no problem, 5 - as large as you can imagine
MM & PJ Jastreboff, 1999

Tinnitus Handicap Inventory (THI)

This form is for informational purposes only and should not take the place of consultation and evaluation by a healthcare professional.

Your Name: Date:

Instructions: The purpose of this questionnaire is to identify, quantify, and evaluate the difficulties that you may be experiencing because of tinnitus. Please do not skip any questions. When you have answer all the questions, add up your total score, based on the values for each response.

1.	Because of your tinnitus, is it difficult for you to concentrate?	Yes (4)	Sometimes (2)	No (0)
2.	Does the loudness of your tinnitus make it difficult for you to hear people?	Yes (4)	Sometimes (2)	No (0)
3.	Does your tinnitus make you angry?	Yes (4)	Sometimes (2)	No (0)
4.	Does your tinnitus make you feel confused?	Yes (4)	Sometimes (2)	No (0)
5.	Because of your tinnitus, do you feel desperate?	Yes (4)	Sometimes (2)	No (0)
6.	Do you complain a great deal about your tinnitus?	Yes (4)	Sometimes (2)	No (0)
7.	Because of your tinnitus, do you have trouble falling to sleep at night?	Yes (4)	Sometimes (2)	No (0)
8.	Do you feel as though you cannot escape your tinnitus?	Yes (4)	Sometimes (2)	No (0)
9.	Does your tinnitus interfere with your ability to enjoy your social activities (such as going out to dinner, to the movies)?	Yes (4)	Sometimes (2)	No (0)
10.	Because of your tinnitus, do you feel frustrated?	Yes (4)	Sometimes (2)	No (0)
11.	Because of your tinnitus, do you feel that you have a terrible disease?	Yes (4)	Sometimes (2)	No (0)
12.	Does your tinnitus make it difficult for you to enjoy life?	Yes (4)	Sometimes (2)	No (0)
13.	Does your tinnitus interfere with your job or household responsibilities?	Yes (4)	Sometimes (2)	No (0)
14.	Because of your tinnitus, do you find that you are often irritable?	Yes (4)	Sometimes (2)	No (0)
15.	Because of your tinnitus, is it difficult for you to read?	Yes (4)	Sometimes (2)	No (0)
16.	Does your tinnitus make you upset?	Yes (4)	Sometimes (2)	No (0)
17.	Do you feel that your tinnitus problem has placed stress on your relationships with members of your family and friends?	Yes (4)	Sometimes (2)	No (0)
18.	Do you find it difficult to focus your attention away from your tinnitus and on other things?	Yes (4)	Sometimes (2)	No (0)
19.	Do you feel that you have no control over your tinnitus?	Yes (4)	Sometimes (2)	No (0)
20.	Because of your tinnitus, do you often feel tired?	Yes (4)	Sometimes (2)	No (0)
21.	Because of your tinnitus, do you feel depressed?	Yes (4)	Sometimes (2)	No (0)
22.	Does your tinnitus make you feel anxious?	Yes (4)	Sometimes (2)	No (0)
23.	Do you feel that you can no longer cope with your tinnitus?	Yes (4)	Sometimes (2)	No (0)
24.	Does your tinnitus get worse when you are under stress?	Yes (4)	Sometimes (2)	No (0)
25.	Does your tinnitus make you feel insecure?	Yes (4)	Sometimes (2)	No (0)

The sum of all responses is your THI Score >>> 0

0-16: Slight or no handicap (Grade 1)
18-36: Mild handicap (Grade 2)
38-56: Moderate handicap (Grade 3)
58-76: Severe handicap (Grade 4)
78-100: Catastrophic handicap (Grade 5)

Newman CW, Jacobson GP, Spitzer JB. (1996) "Development of the Tinnitus Handicap Inventory." Archives of Otolaryngology - Head and Neck Surgery. 122(2):143-8.
McCombe, A., Baguley, D., Coles, R., McKenna, L., McKinney, C. & Windle-Taylor, P. (2001). "Guidelines for the Grading of Tinnitus Severity: the Results of a Working Group Commissioned by the British Association of Otolaryngologists, Head and Neck Surgeons." Clinical Otolaryngology, 26, 388-393.

Agenda

Categorization in TRT

Category	Hyperacusis	Prolonged sound-induced exacerbation	Subjective hearing loss	Impact on life	Treatment
0	-	-	-	low	Counseling only
1	-	-	-	high	Sound generator set at mixing point
2	-	-	present	high	Hearing aid with stress on enrichment of the auditory background
3	present	-	not relevant	high	Sound generators set above threshold of hearing
4	present	present	not relevant	high	Sound generators set at threshold; very slow increase of sound level

Decision table

- **Objects**: patient visits (initial)
- **Conditions**: demographics, medical evaluation, patient's interview,. . .
- **Decision**: category of a hearing problem (as defined by TRT)

THC#	Demographics				Etiology	Onset	Initial Interview			Audiometry			Medications	Category
ID	Age	Gender	Occup	...			Tinnitus	Decreased Sound Tolerance	Hearing Loss	Pure tone audiogram	Loudness Discomfort Levels	...		
1														
2														

Decision rules: Interview → Category

- Database:

THC#	Demographics				Etiology	Onset	Initial Interview			Audiometry			Medications	Category
ID	Age	Gender	Occup	...			Tinnitus	Decreased Sound Tolerance	Hearing Loss	Pure tone audiogram	Loudness Discomfort Levels	...		
1							7	0	0					1

- Example:

IF H_pr(<0;0.5)) AND HL_pr(<0;0.5)) AND T_pr(<6;8))
THEN Category(1), Conf.= 85%

- Interpretation: *If* hyperacusis is not a problem (*H_pr* score less than 0.5) and hearing loss is not a problem (*HL_pr* score less than 0.5), and tinnitus indicated as a problem *T_pr* with scores in range <6;8), *then* a patient is categorized into Category 1, with 85% probability.

Decision rules: Audiometry → Category

- Database:

THC#	Demographics				Etiology	Onset	Initial Interview			Audiometry		Med.	Category
ID	Age	Gender	Occup	...			Tinnitus	Decreased Sound Tolerance	Hearing Loss	Pure tone audiogram	Loudness Discomfort Levels		
1										R2=45; R6=75	LR8=999		2

- Example:

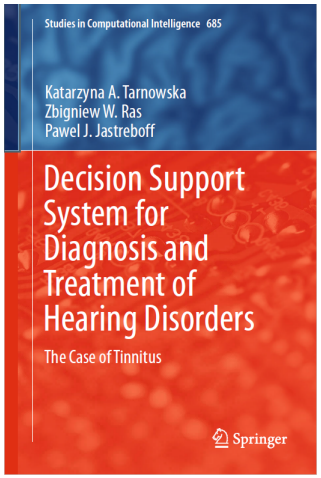
IF LR8>=999 AND R2>=45 AND R6>=75 THEN Category(2), Conf.= 89%

- Interpretation: If LDL in norm and a pure-tone audiogram with 2kHz of right ear test is equal or greater than 45, and with the pure-tone threshold of 6kHz is equal or greater than 75, then a patient is categorized into Category 2, with 89% probability.

Diagnostic decision rules for TRT - reference

More examples of diagnostic decision rules in:

- Chapter 6. Experiment 2: Diagnostic Rules



Agenda

Agenda

LERS - **L**earning from **E**xamples based on **R**ough **S**ets

- The first implementation of LERS was done by John S. Dean and Douglas J. Sikora in 1988
- Input data is represented as a *decision table*.
- In the decision table, examples are described by values of *attributes* and characterized by a value of a *decision*.
- All examples with the same value of the decision belong to the same *concept*.
- This algorithm looks for regularities in the decision table.

LERS - example information system

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

LERS - example classification attributes

Let $S = (X, A, V)$ be the information system.

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Classification attributes

LERS - example decision attribute

Let $S = (X, A, V)$ be the information system.

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Decision attribute

LERS LEM1 algorithm - example

Step 1: compute partitions generated by a single attribute

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

The partitions of *X*, generated by single attributes are:

$$\{a\}^* = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$$

$$\{b\}^* = \{\{x_1, x_3\}, \{x_2, x_4, x_5, x_6\}, \{x_7, x_8\}\}$$

$$\{c\}^* = \{\{x_1, x_3, x_5, x_6\}, \{x_2, x_4\}, \{x_7, x_8\}\}$$

$$\{d\}^* = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8\}\}$$

Let C be the set containing of one attribute {f}:

$$\{f\}^* = \{\{x_1, x_3\}, \{x_2, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$$

LERS LEM1 algorithm - example

Step 2: Check if partitions generated by conditional attributes are subsets of partition generated by the decision attribute

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

The partitions of *X*, generated by single attributes are:

$$\{a\}^* = \{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$$

$$\{b\}^* = \{\{x_1, x_3\}, \{x_2, x_4, x_5, x_6\}, \{x_7, x_8\}\}$$

$$\{c\}^* = \{\{x_1, x_3, x_5, x_6\}, \{x_2, x_4\}, \{x_7, x_8\}\}$$

$$\{d\}^* = \{\{x_1, x_2, x_3, x_4, x_5, x_6\}, \{x_7, x_8\}\}$$

Let C be the set containing of one attribute {f}:

$$\{f\}^* = \{\{x_1, x_3\}, \{x_2, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$$

None of the sets is a subset of {f}*

LERS LEM1 algorithm - example

Step 3: Compute partitions for two-item sets. Check subset relations.

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

forming two item sets:

$$\{a,b\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{a,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{a,d\}^* = \{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} = \{a\}^*$$

$$\{b,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{b,d\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{c,d\}^* = \{\{x_1,x_3,x_5,x_6\},\{x_2,x_4\},\{x_7,x_8\}\} = \{c\}^*$$

$$\{f\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\}$$

LERS LEM1 algorithm - example

Step 4: Mark partitions that are subsets (covered) of the decision's partition

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

$\{a,b\}* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}*$

$\{a,c\}* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}*$

$\{a,d\}* = \{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} = \{a\}*$

$\{b,c\}* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}*$

$\{b,d\}* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}*$

$\{c,d\}* = \{\{x_1,x_3,x_5,x_6\},\{x_2,x_4\},\{x_7,x_8\}\} = \{c\}*$

marked

$$\{f\}* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\}$$

LERS LEM1 algorithm - example

Step 5: Mark partitions that are subsets (covered) of one-item sets

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

$$\{a,b\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{a,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{a,d\}^* = \{\{x_1,x_2,x_3,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} = \{a\}^*$$

$$\{b,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{b,d\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{c,d\}^* = \{\{x_1,x_3,x_5,x_6\},\{x_2,x_4\},\{x_7,x_8\}\} = \{c\}^*$$

marked, but not covering of f

$$\{f\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\}$$

LERS LEM1 algorithm - example

Step 6: Check if partitions all are marked (if not repeat 3-5 for three-item sets). Check coverings for the decision (coverings of C)

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

$$\{a,b\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$
$$\{a,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

$$\{b,c\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$
$$\{b,d\}^* = \{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_8\}\} \subseteq \{f\}^*$$

All of the sets are marked!

The coverings of C are:

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

LERS LEM1 - how to find rules from coverings?

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Covering {a,b}

$(a,0)^* = \{x_1, x_2, x_3, x_4\}$

$(a,1)^* = \{x_5, x_6\} \subseteq (f,2)^*$

$(a,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$(b,0)^* = \{x_1, x_3\} \subseteq (f,0)^*$

$(b,1)^* = \{x_2, x_4, x_5, x_6\}$

$(b,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

marked

$\{f\}^* = \{\{x_1, x_3\}, \{x_2, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$

LERS LEM1 - how to find rules from coverings?

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Covering {a,b}

$(a,0)^* = \{x_1, x_2, x_3, x_4\}$

$(a,1)^* = \{x_5, x_6\} \subseteq (f,2)^*$

$(a,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$(b,0)^* = \{x_1, x_3\} \subseteq (f,0)^*$

$(b,1)^* = \{x_2, x_4, x_5, x_6\}$

$(b,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$((a,0) \wedge (b,1))^* = \{x_2, x_4\} \subseteq (f,1)^*$

marked

$\{f\}^* = \{\{x_1, x_3\}, \{x_2, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$

LERS LEM1 - how to find rules from coverings?

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Covering {a,b}

Certain rules, obtained from marked items:

$$(a,1) \rightarrow (f,2)$$
$$(a,2) \rightarrow (f,3)$$
$$(b,0) \rightarrow (f,0)$$
$$(b,2) \rightarrow (f,3)$$
$$(a,0) \wedge (b,1) \rightarrow (f,1)$$

$(a,1)^* = \{x_5, x_6\} \subseteq (f,2)^*$

$(a,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$(b,0)^* = \{x_1, x_3\} \subseteq (f,0)^*$

$(b,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$((a,0) \wedge (b,1))^* = \{x_2, x_4\} \subseteq (f,1)^*$

LERS LEM1 - how to find rules from coverings?

Let $S = (X, A, V)$ be the information system.

X	a	b	c	d	f
x1	0	0	0	1	0
x2	0	1	1	1	1
x3	0	0	0	1	0
x4	0	1	1	1	1
x5	1	1	0	1	2
x6	1	1	0	1	2
x7	2	2	2	0	3
x8	2	2	2	0	3

Covering {a,b}

Possible rules, obtained from non-marked items:

$(a,0) \rightarrow (f,0)$ with confidence ½

$(a,0) \rightarrow (f,1)$ with confidence ½

$(b,1) \rightarrow (f,1)$ with confidence ½

$(b,1) \rightarrow (f,2)$ with confidence ½

$(a,1)^* = \{x_5, x_6\} \subseteq (f,2)^*$

$(a,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$(b,0)^* = \{x_1, x_3\} \subseteq (f,0)^*$

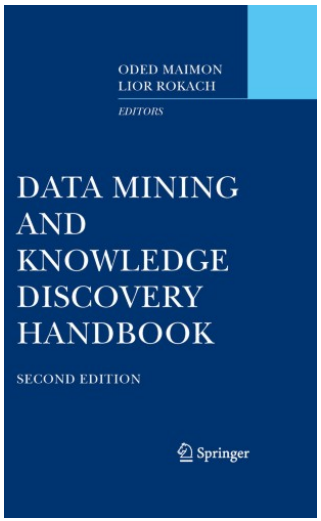
$(b,2)^* = \{x_7, x_8\} \subseteq (f,3)^*$

$((a,0) \wedge (b,1))^* = \{x_2, x_4\} \subseteq (f,1)^*$

LERS (LEM1) references

For a complete description of the algorithm on rule induction presented (LEM1) see:

- Grzymala-Busse J.W. (2009) Rule Induction. In: Maimon O., Rokach L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA.
- Available at `https://link.springer.com/chapter/10.1007/978-0-387-09823-4_13`, listed as optional reading under the current week on Canvas



Agenda

Graded Lab

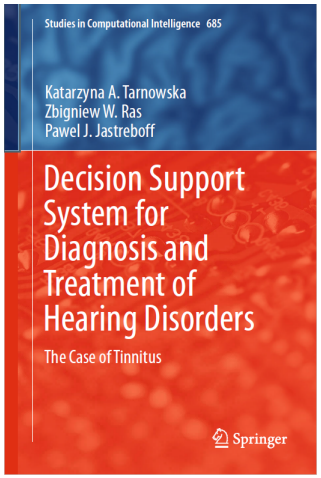
Lab 3 activity: data pre-processing

Working with the dataset you chose in Lab 1/2 and answer the following questions:

- Describe any incomplete, noisy, or inconsistent data found in your dataset.
- Apply at least one pre-processing technique in WEKA (i.e. discretization)
- Perform attribute selection, using and comparing at least three methods implemented in WEKA.

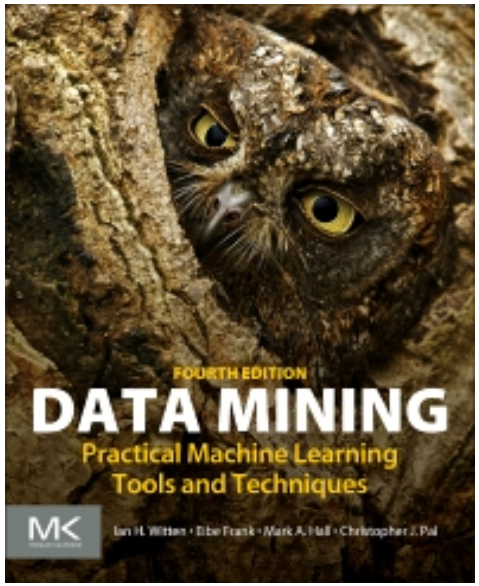
Lab 3 - Sample case study for tinnitus data

- Chapter 6.1: Initial Feature Development
- Chapter 6.2.2: Feature Selection
- Chapter 6.3.1-6.3.2: Pharmacology Data Analysis; Pivotal Features Development



Lab 3 - reference

- Chapter 8: Data transformations



Self-check

Make sure you know:

1. The definition of the *decision rule*.
2. The definition of *support* and *confidence*.
3. An example scenario for decision rules application.
4. How to apply the LERS LEM1 algorithm for rule induction in a sample information system.