

---

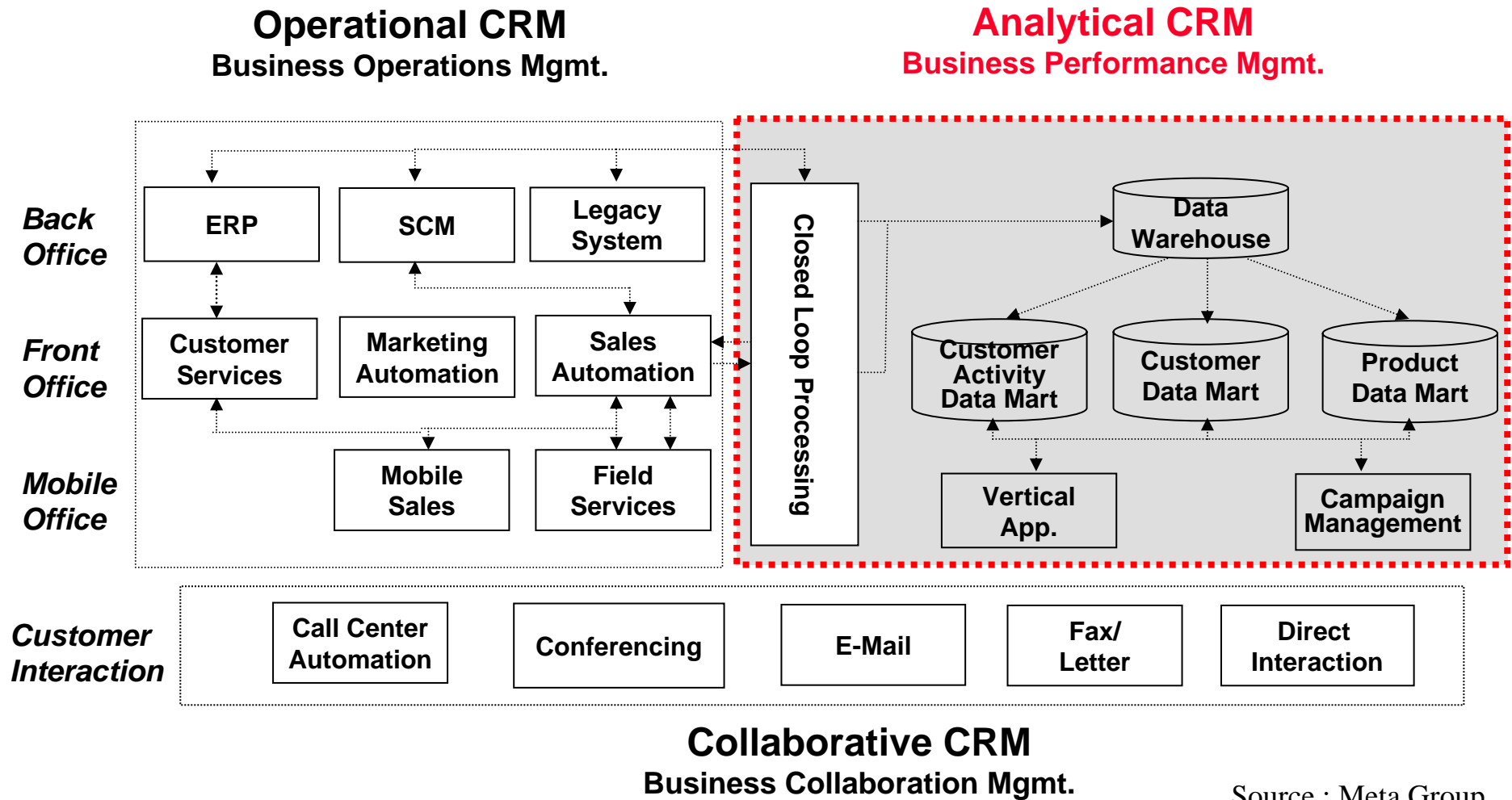
# Analytic CRM

2006. 5. 11

*tsshin@yonsei.ac.kr*

- 
- Analytic CRM
  - Analytic CRM
  - Data Mining

# Analytical CRM in CRM Ecosystem



Source : Meta Group

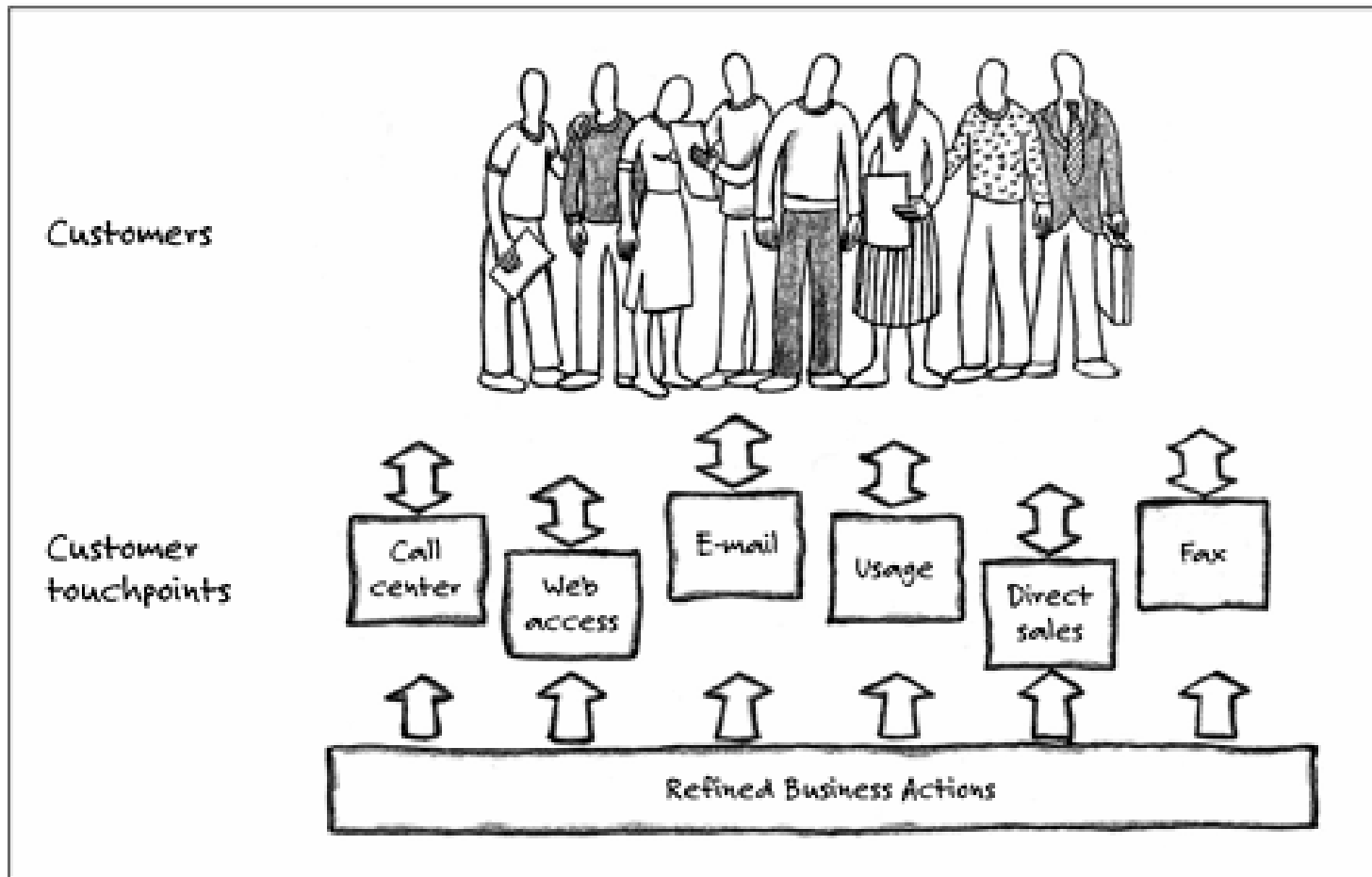
# Analytic CRM

---

## ■ Analytic CRM

- The capture, storage, extraction, processing, interpretation, and reporting of customer data to a user.
- Analytical CRM, also known as "back-office" or "strategic" CRM, involves understanding the customer activities that occurred in the front office.
- Analytical CRM requires technology (to compile and process the mountains of customer data to facilitate analysis) and new business processes (to refine customer-facing practices to increase loyalty and profitability).
- Under pressure from analysts and industry experts, most of today's CRM vendors are either creating analytical CRM capabilities or partnering with business intelligence (BI) vendors to incorporate analysis into their offerings.

# Operational CRM: Touching the customer (Figure 1-1)



(Dyche, 2002)

# Analytical CRM: Understanding the customer (Figure 1-2)

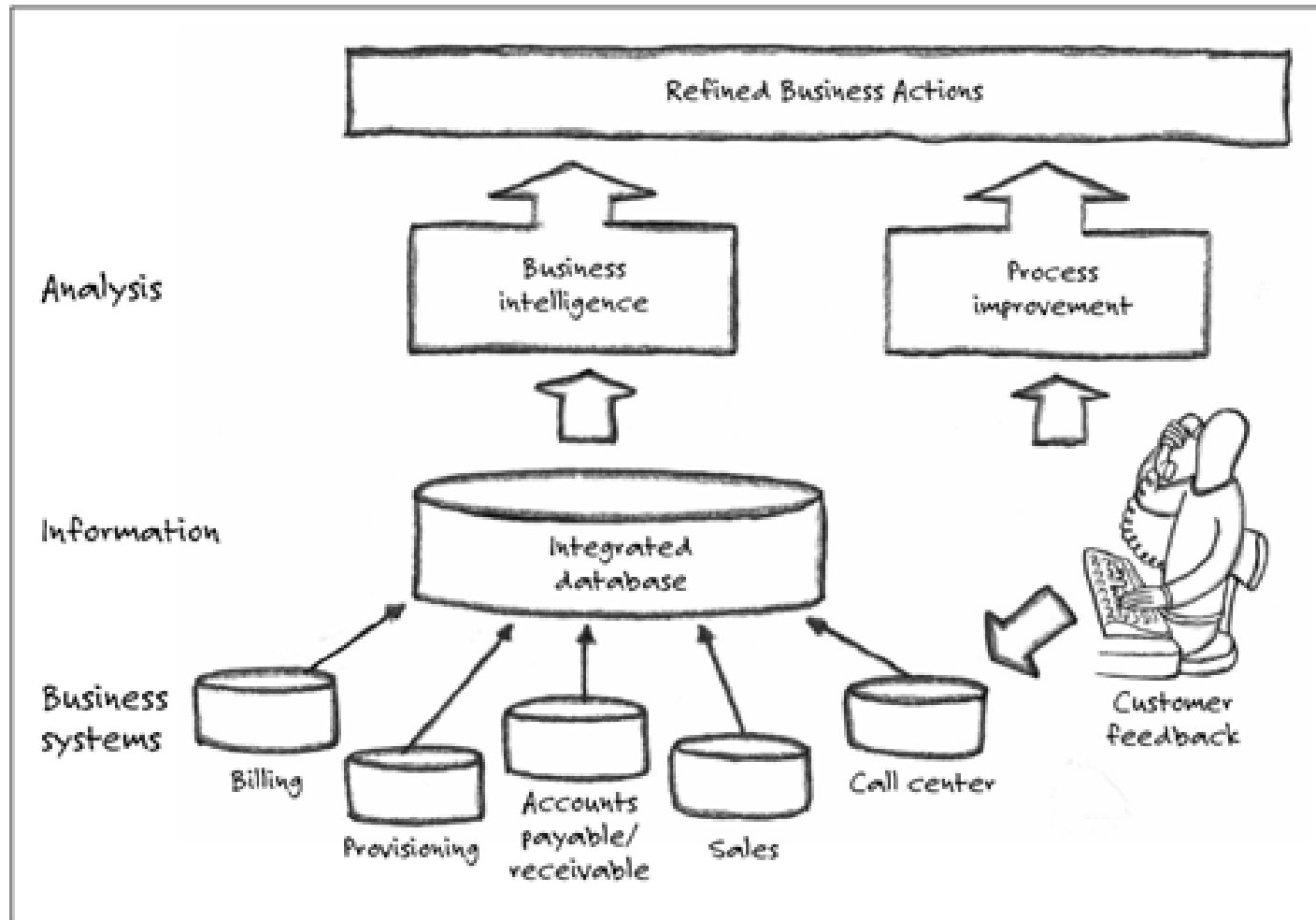
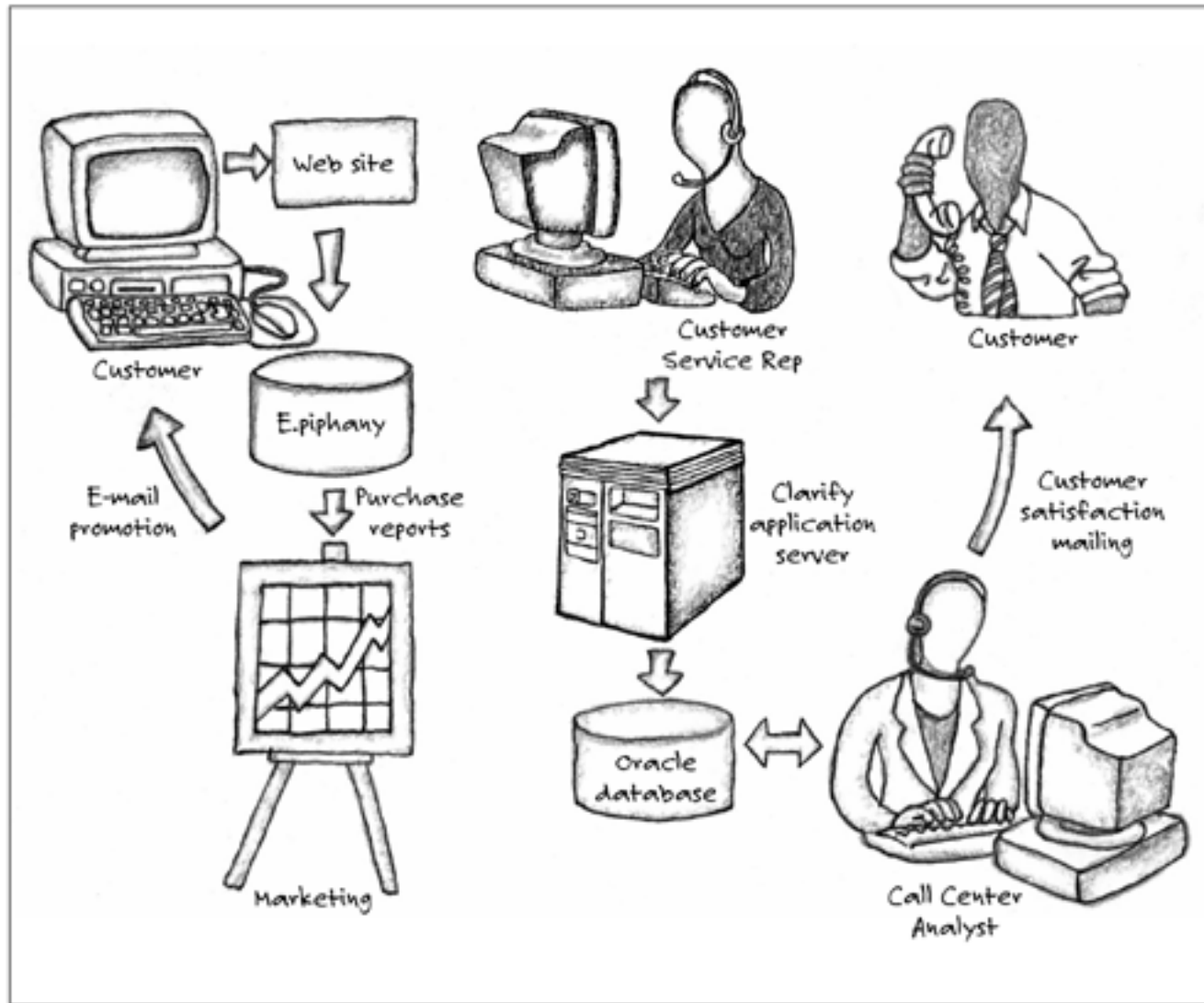


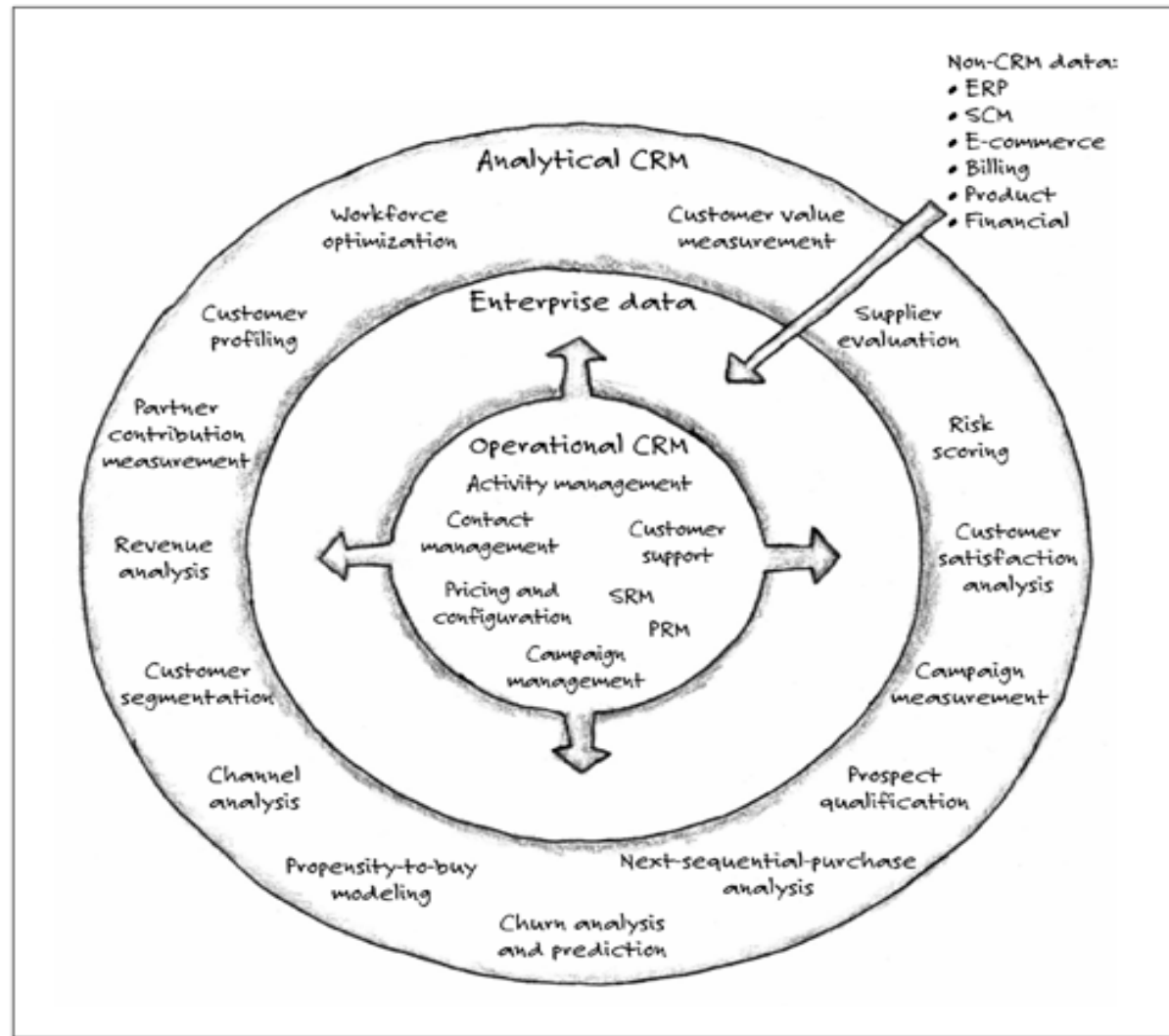
Figure 1-2 shows how the data and processes combine to refine business actions. (Dyche, 2002)

# One company, two CRM systems (Figure 6-1)



(Dyche, 2002)

# Analytical CRM: The sum of its parts (Figure 6-4)



(Dyche, 2002)



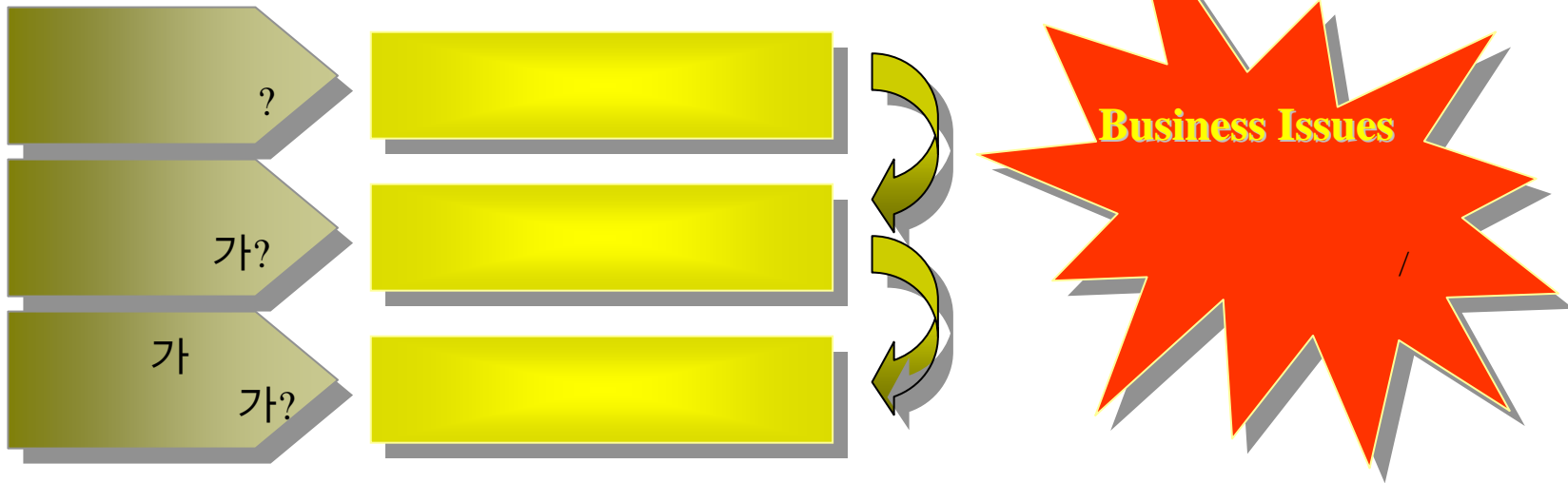
# Analytic CRM

---

- 1: A Churn Management
- 2: B DB Marketing
- 3: DM (Microsoft)
- 4: Market-basket Analysis ( )

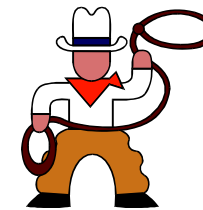
# 1: A

# Churn Management



## ■ Churn Management

- - 
  - 
  - ( )
  - , 가
- - 
  -

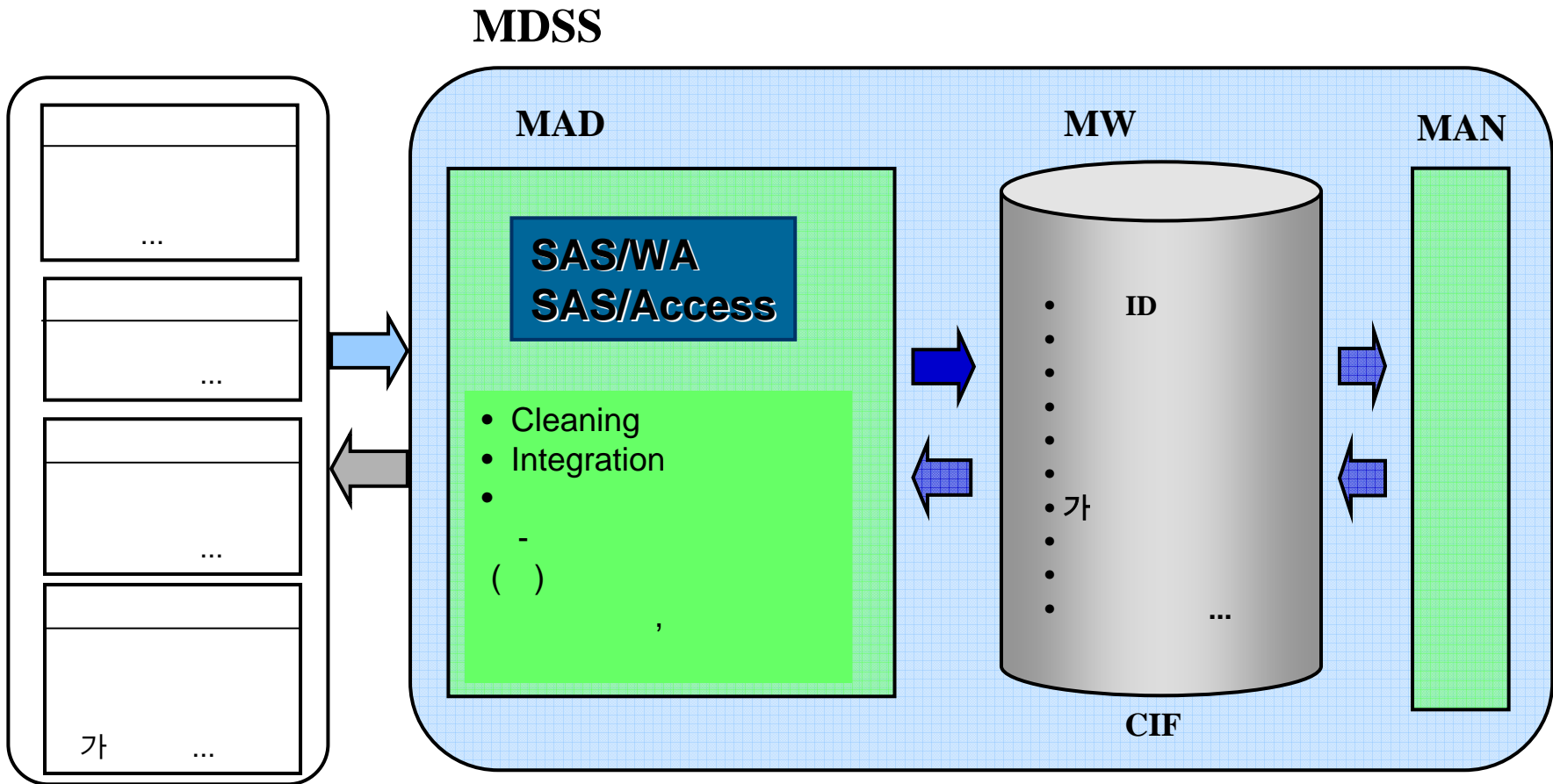


# 1: A

# Churn Management

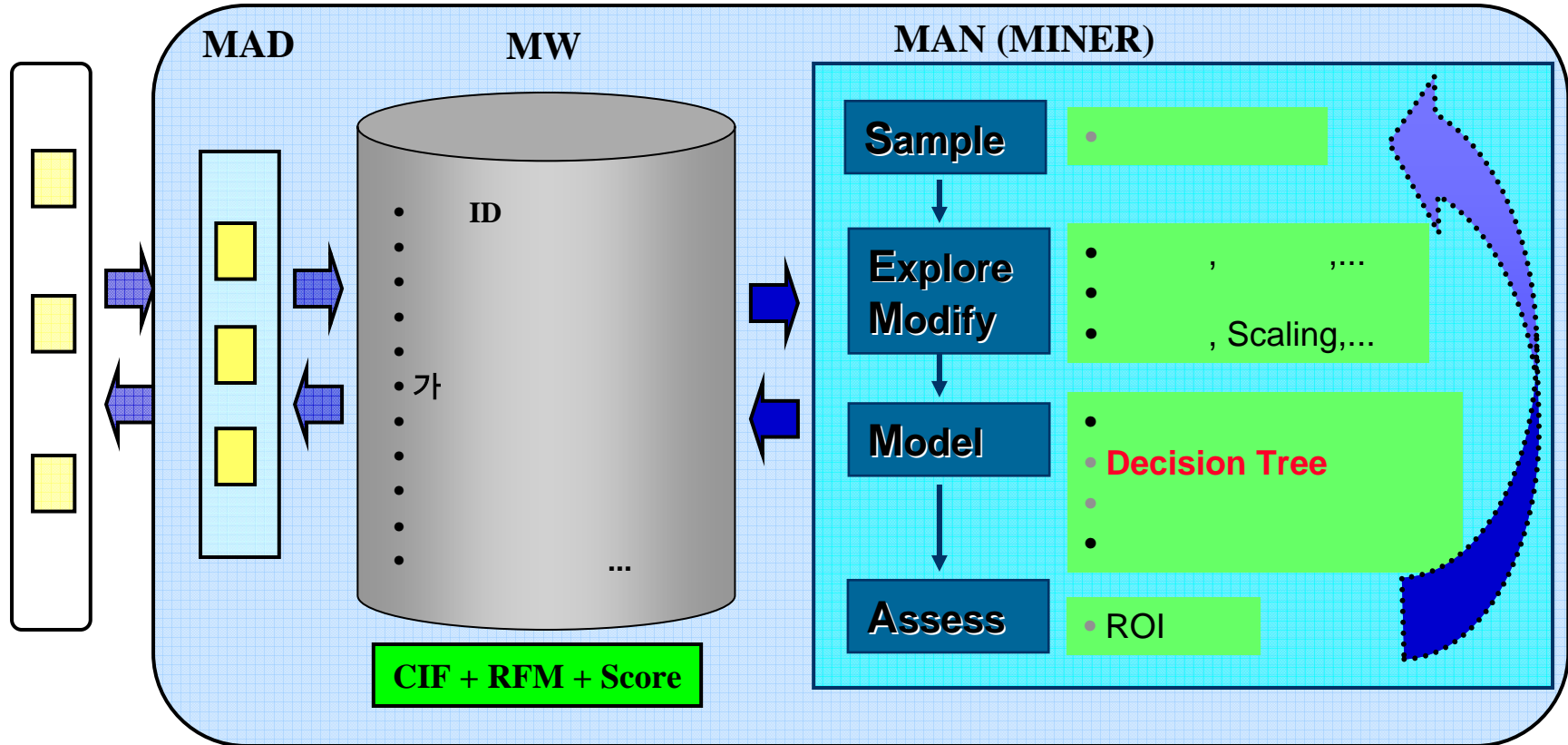
( )

■ MW(Marketing Warehouse) : MAD



■ : MAN (MINER)

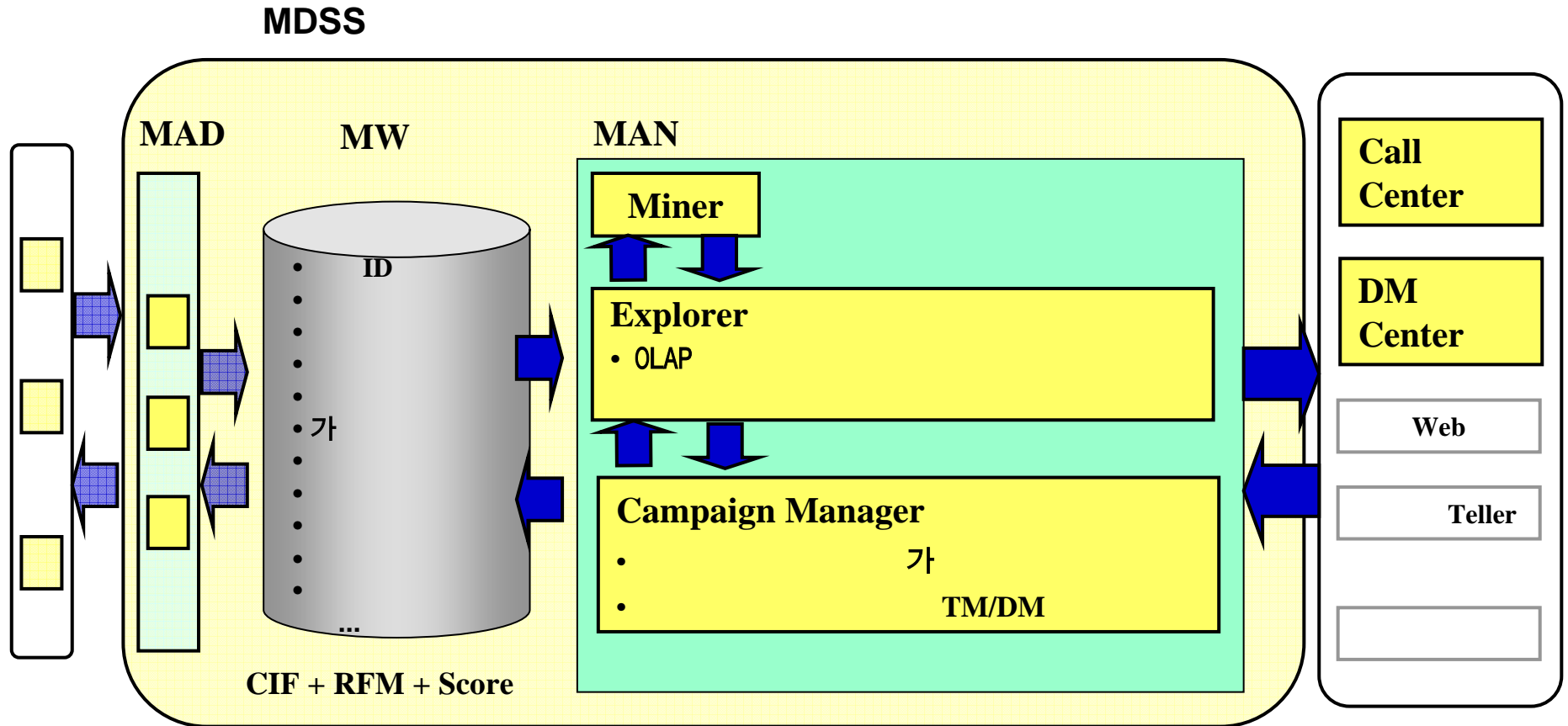
## MDSS



# 1: A

# Churn Management ( )

■ : MAN



## 2: B

## DB Marketing

1.

1)

—

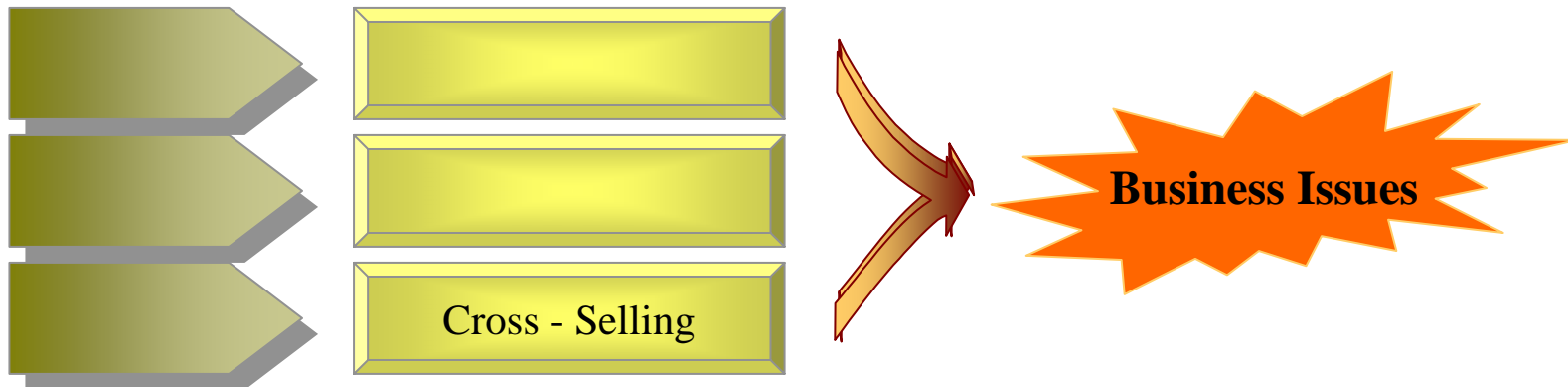
— Cross-Selling

2)

—

, TM , DM

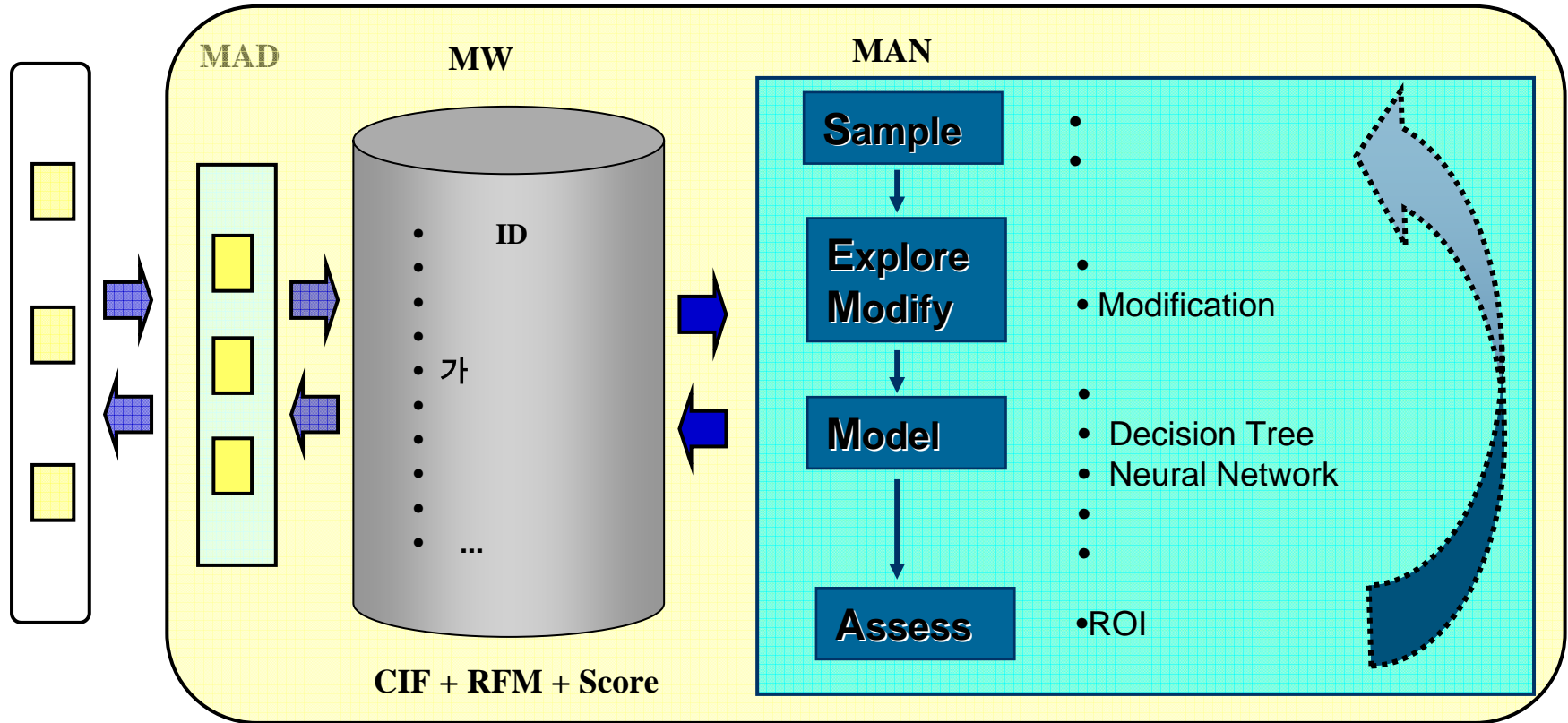
3)



2.

: MAN

MDSS





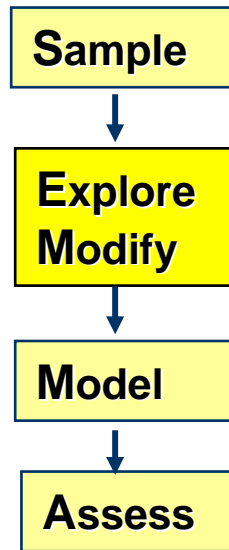
## 2: B

## DB Marketing

( )

3. ( )

- 
- 
- : MAN



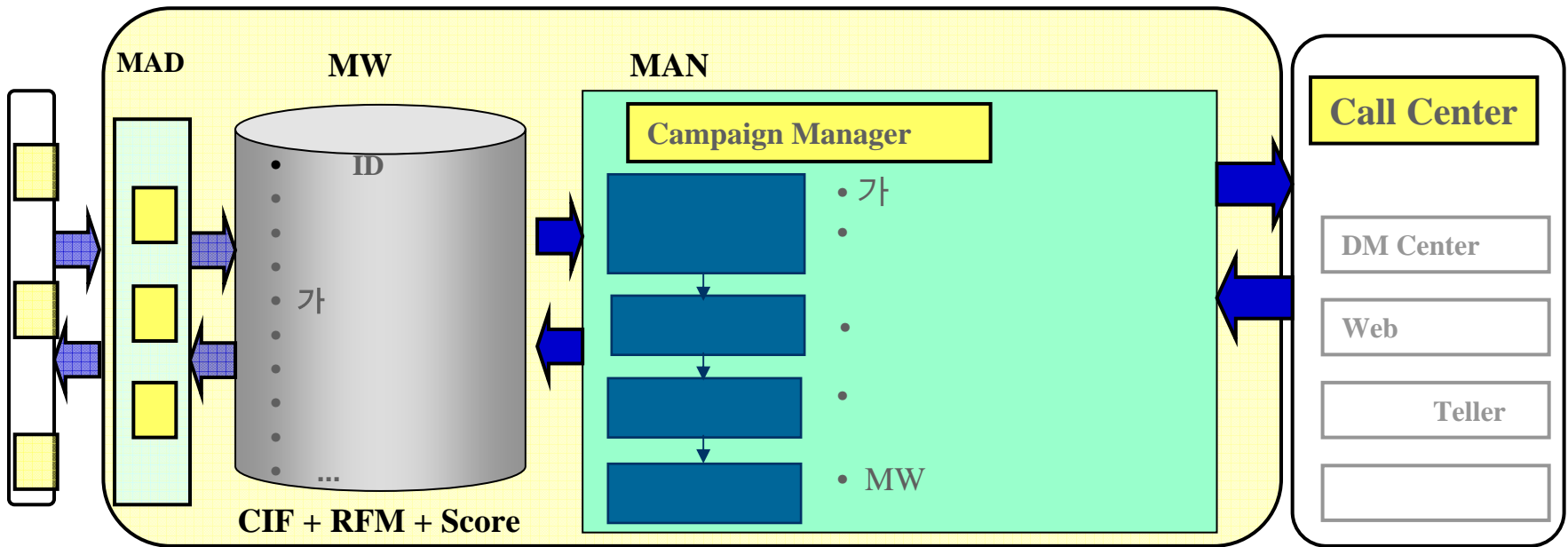
- 
- ( )
- 
- 
- 
- 
- 

가/  
**Channel**

# 2: B DB Marketing ( )

## 4. ( ): Step 1

- 
- 
- Campaign Manager : MAN



## 2: B DB Marketing ( )

5. ( ) : Step 2

- 
- 
- : MAN

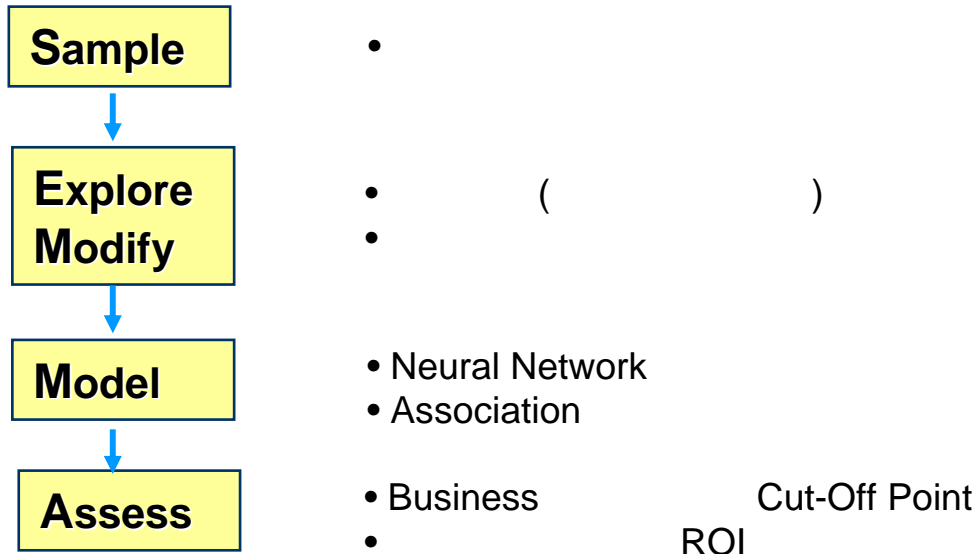


# 2: B DB Marketing ( )

## 6. Cross - Selling( )

-

- 2 가



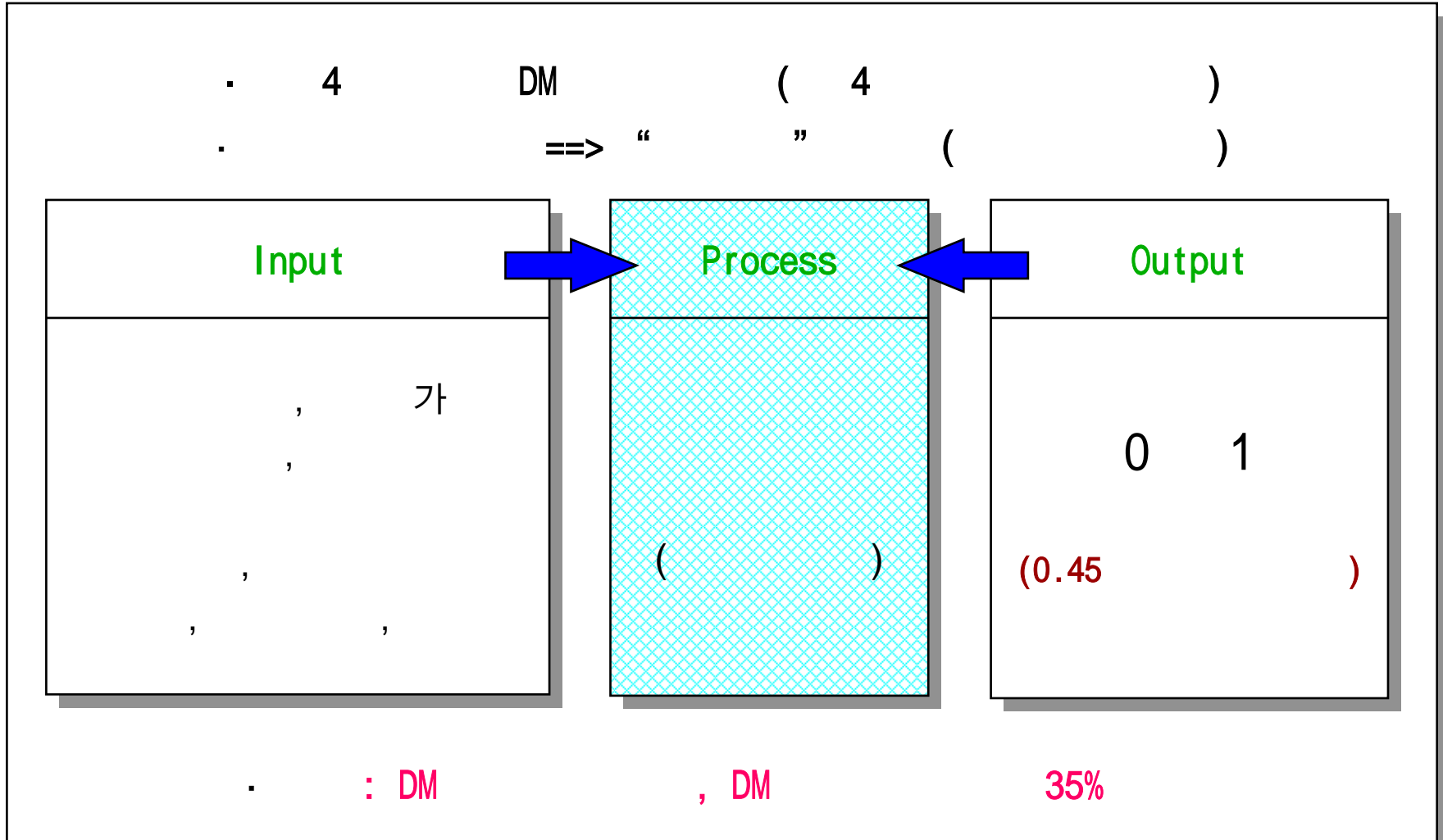
# 2: B DB Marketing ( )

7.

- 1) ( )
  - : , Call Center, DM
  - DM, TM
  - 가 3% ( 2~3 )가
  
- 2) ( )
  - 가 가
- 3) Cross-selling ( )
  - 가 40%
  - 가 가 20%
  - 가

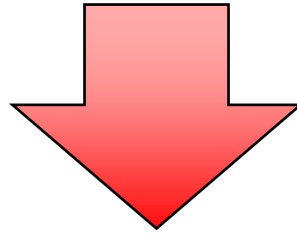
# 3: DM

# (Microsoft)



# 4: Market-basket Analysis ( )

○ : “ 1 가? ”



○ ⇒ “ 가 가 , ? ”

---

# Data Mining



# Data Mining

---

- Data Mining
- Data Mining
- Data Mining
- Data Mining
- Data Mining
  - 
  - Rule Induction
  - SONN
  -

# Data Mining

---



1960

.



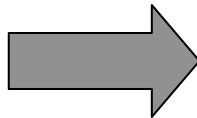
.



,

,

**Data**



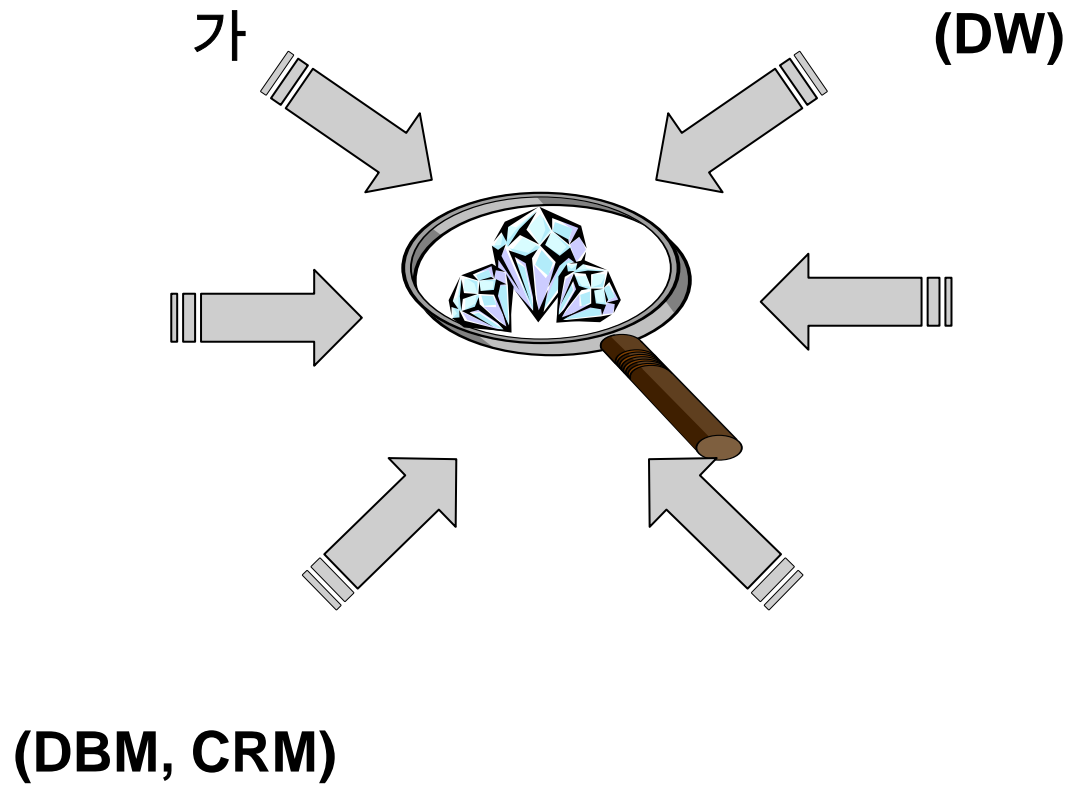
# Data Mining ( )

---



- "Data mining (mathematical) (Statistics) (Neural Networks) 가 (pattern recognition technologies) ." (Gartner Group)
- Data mining (unknown), 가 (actionable information) . ( Aaron Zornes, The META Group )

# Data Mining ( )



# Data Mining

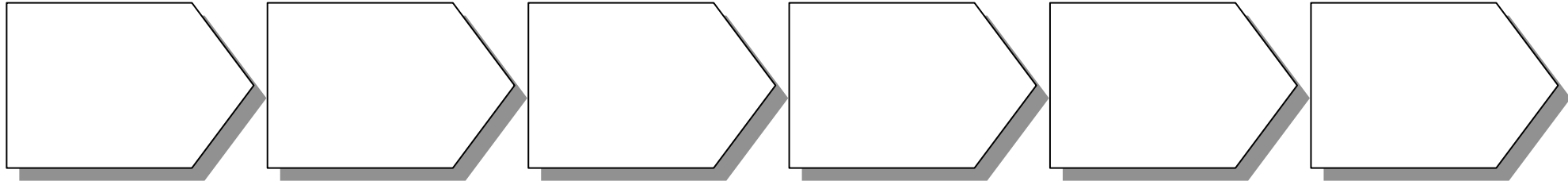
---

- **(Descriptive Modeling)**
  - /
  - ( ) .
  - (Unsupervised learning) .
  - , (Clustering/Segmentation) .
  - , .
- **(Predictive Modeling)**
  - 
  - ( ) .
  - (Supervised learning) .
  - (Classification), (Value Prediction) .
  - , (Rule Induction),  
( , , , Probit ) .

		Predictive		Descriptive	
		Classification	Estimation	Affinity Grouping	Clustering
	ANN	✓	✓		✓ (Kohonen Network)
	CBR	✓	✓		
	Tree Induction	✓			
	Regression	✓	✓		
	Association Rule			✓	
	Clustering				✓

(ANN: Artificial Neural Network, CBR: Case-based Reasoning)

# Data Mining



•  
•

•

•

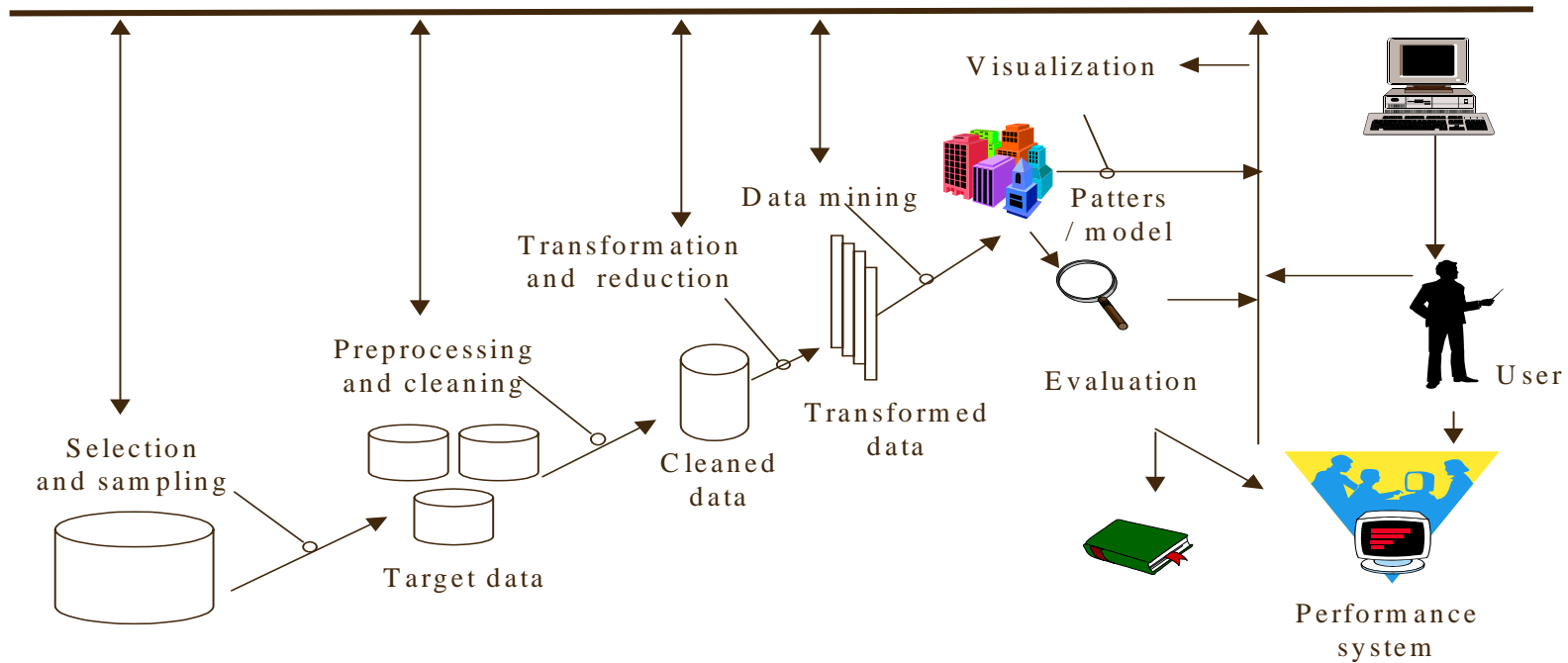
•

•

•

•

Scoring



- ( )
  - .
- , ( )
  - (t- , ANOVA, )
  - ( )
  - (Logistic Regression, Probit, MDA)
  -



# Data Mining

---

- 가?
- 가 가?
- 가 가?
- ( .)
- 가?
- .
- 가?

---

# (Artificial Neural Networks)



- : " ”
- 
- 가
- 가
- .

가

가

.

( )

---



**(Caudill and Butler, 1992)**



,

**(Hecht-Nielsen, 1991)**

---

■ 1940

■ : 1943

(McCulloch) (Pitts)

가



■ (Hebb)

( , neuron)

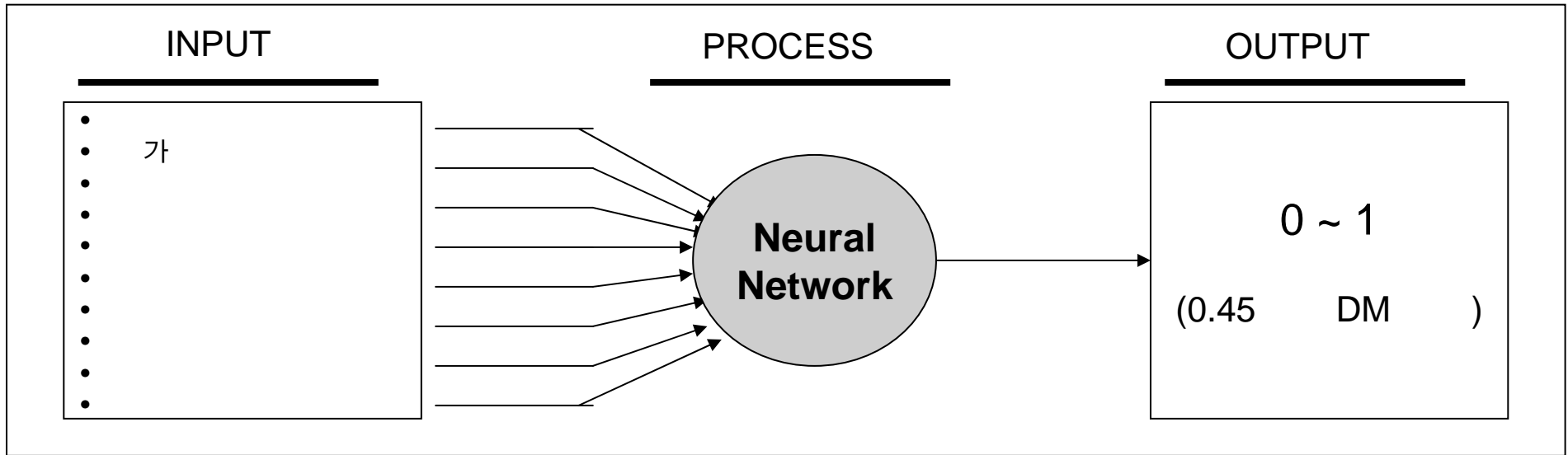
(Weight)

( )

- 1957 (Rosenblatt) (Perceptron)
- Widrow가 Adaline(Adaptive Linear)
  - 
  - 
  - XOR(Exclusive OR)
- 1980 : Hopfield, Rumelhart, McClelland
  - , “ ”
  - 가 (Error Backpropagation)

# (ANN:Artificial Neural Network)

- ,
- 
- Input Layer, Hidden Layer, Output Layer  
Node
- Node Weight( ) , Weight
- Black Box



,



- , .
- ,
- ,
- ,

.  
.

가 가



- 
- 
- 

->

(local minimum)

(Black box)



# XOR(Exclusive OR)

---

■ ( ) 가

.

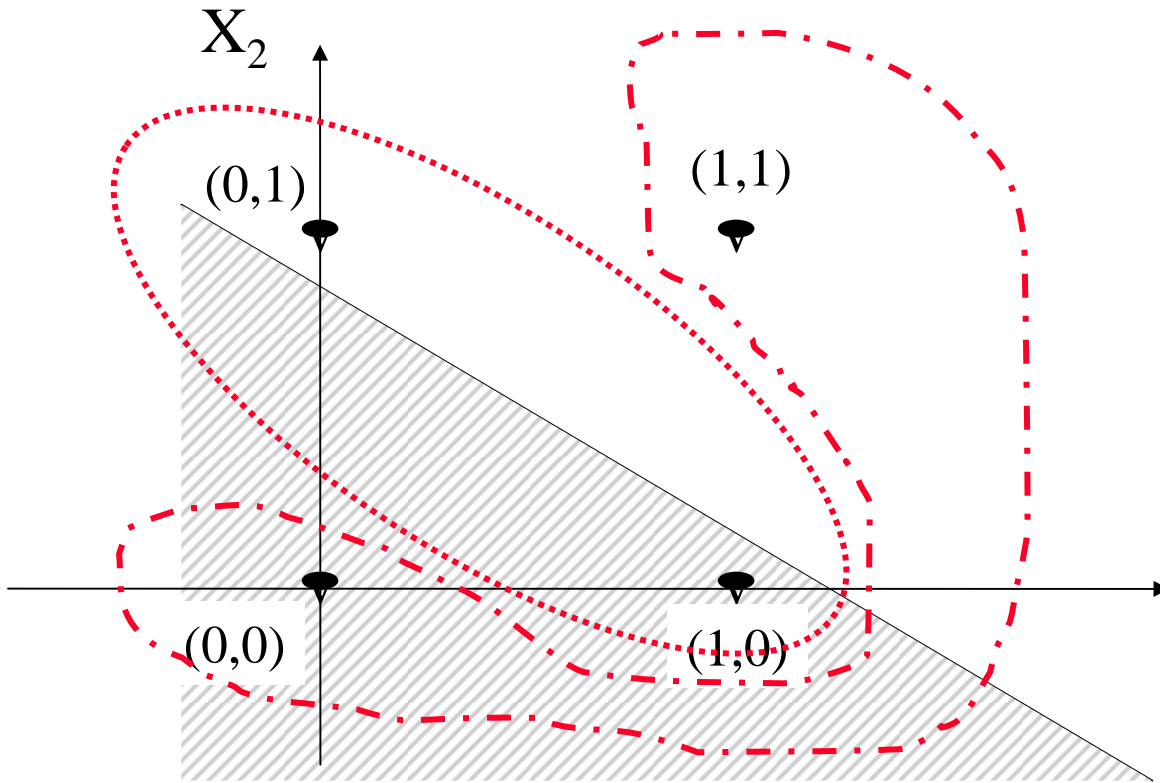
■ XOR OR , OR ,

XOR

■ XOR

가

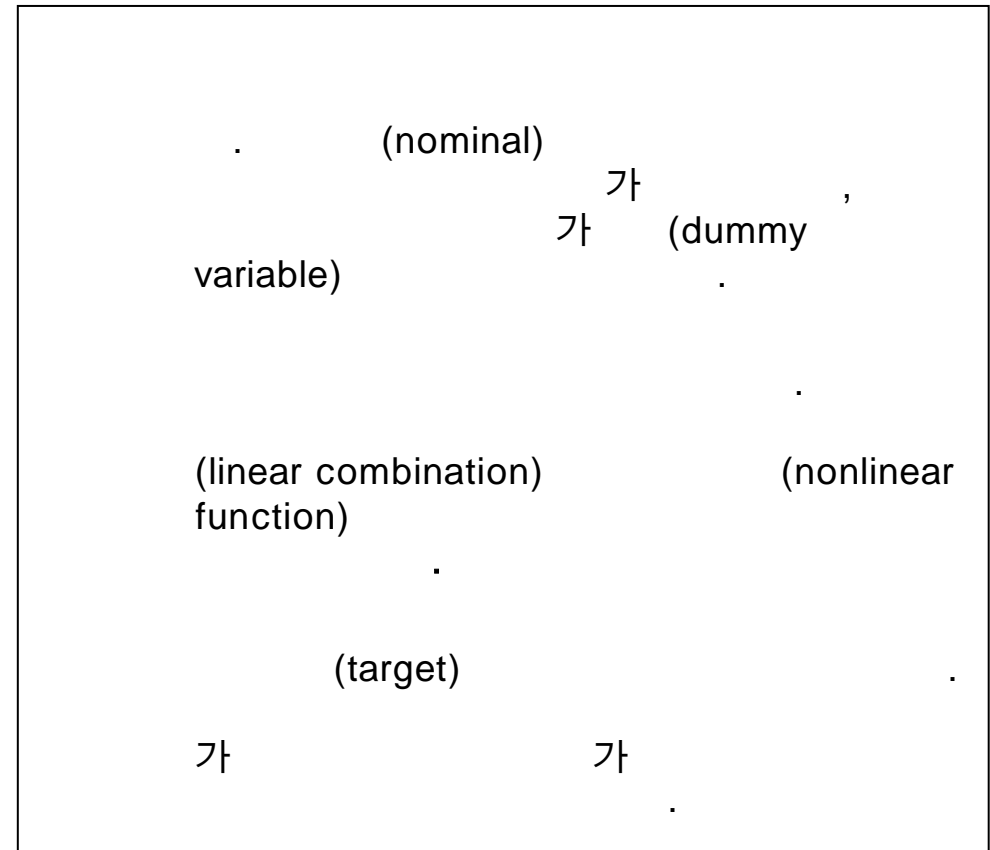
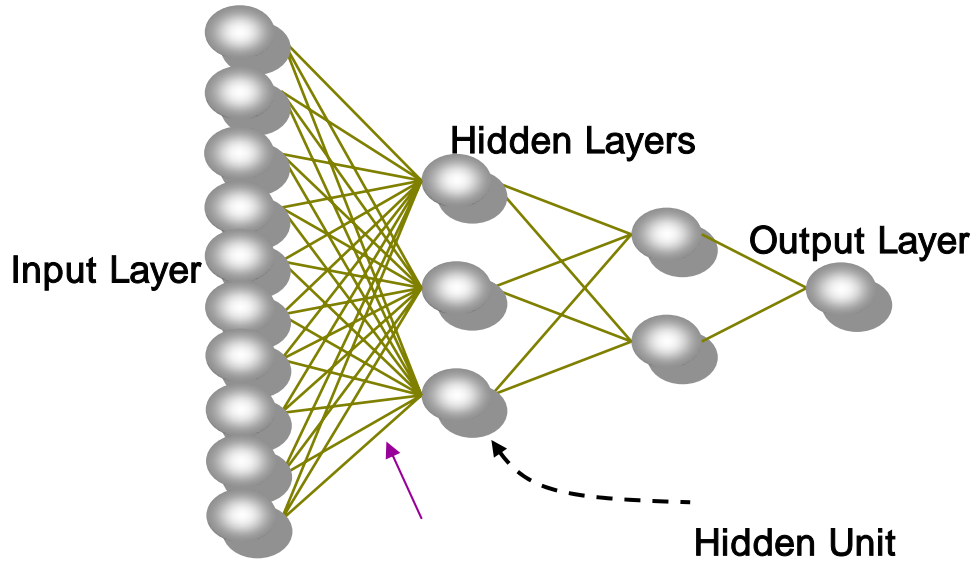
# XOR(Exclusive OR) ( )



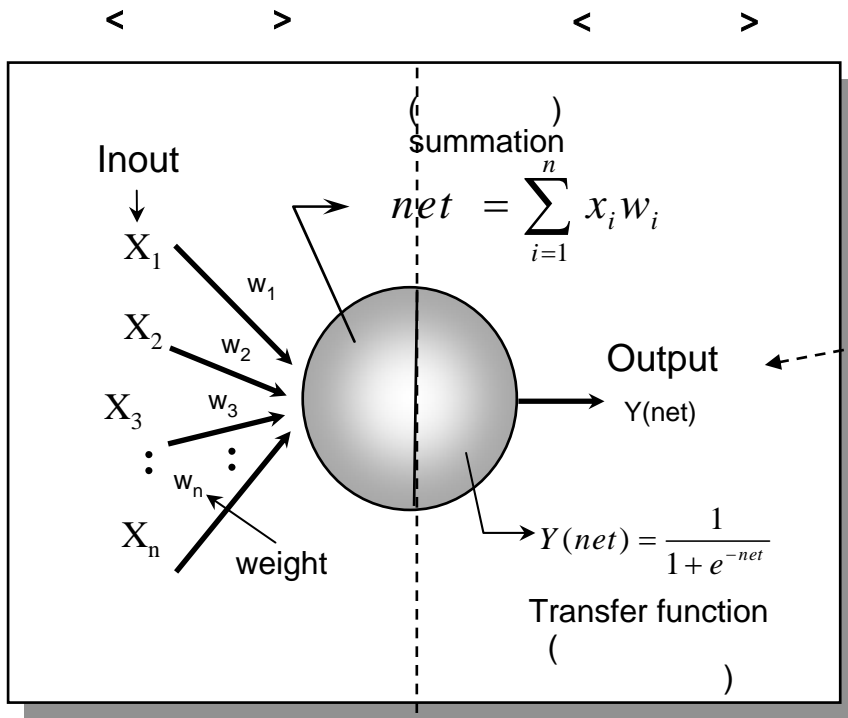
$X_1$	$X_2$	$\rightarrow$	$Y$
1	1		0
1	0		1
0	1		1
0	0		0

$$x_1 w_1 + x_2 w_2 > Threshold$$

- 가 (Node) , 가 Hidden Layer, 가 (Hidden Layer) (Hidden Layer), 가 (Weight) Hidden Layer ) (Hidden Layer)

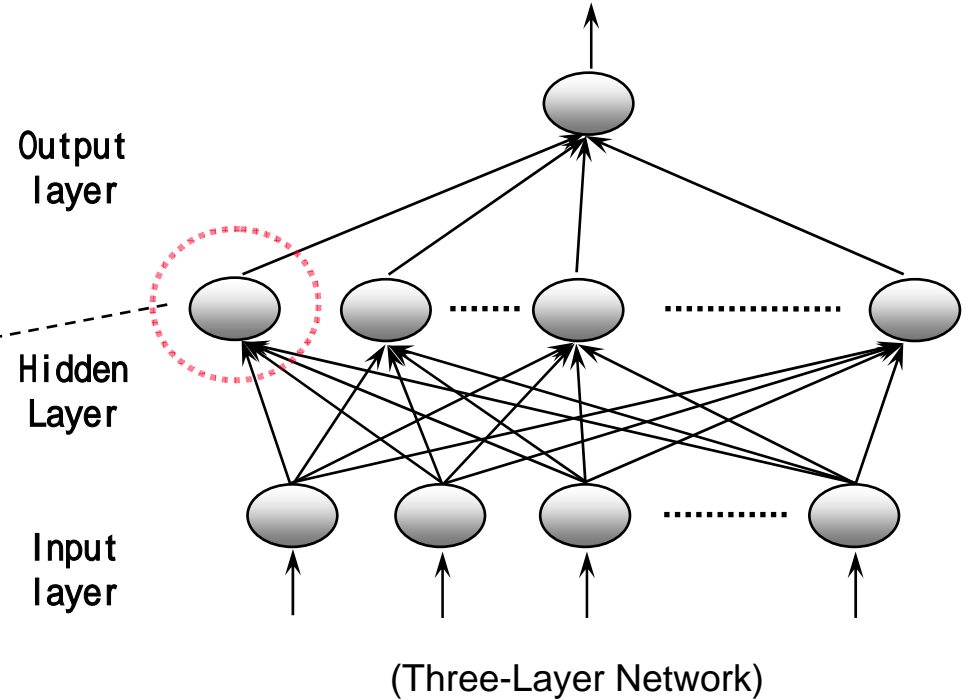


( )



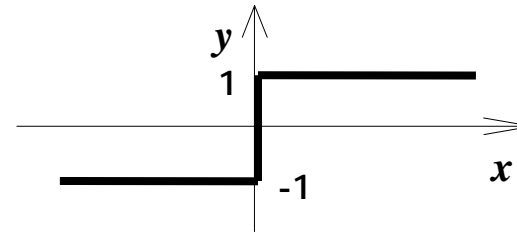
(Node)

(Neuron)

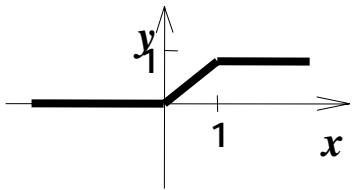


# (Transfer function)

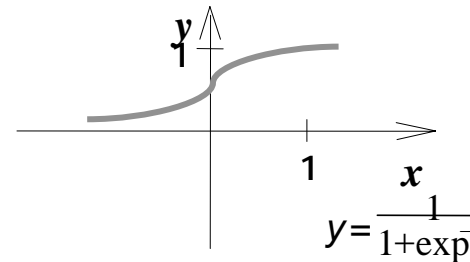
Types of Transfer Function	Equation
Hard Limiter	$x < 0, y = -1$ $x \geq 0, y = 1$
Ramping Function	$0 \leq x < 1, y = x$ $x > 1, y = 1$
Sigmoid Function I	$y = \frac{1}{1 + \exp^{-x}}$
Sigmoid Function II	$x \geq 0, y = 1 - \frac{1}{1 + x}$ $x < 0, y = 1 + \frac{1}{1 - x}$
Hyperbolic Tangent Function	$y = \frac{\exp^x - \exp^{-x}}{\exp^x + \exp^{-x}}$



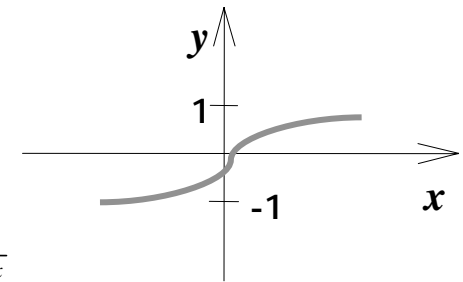
**Hard Limiter**



**Rapping Function**



**Sigmoid Function I**



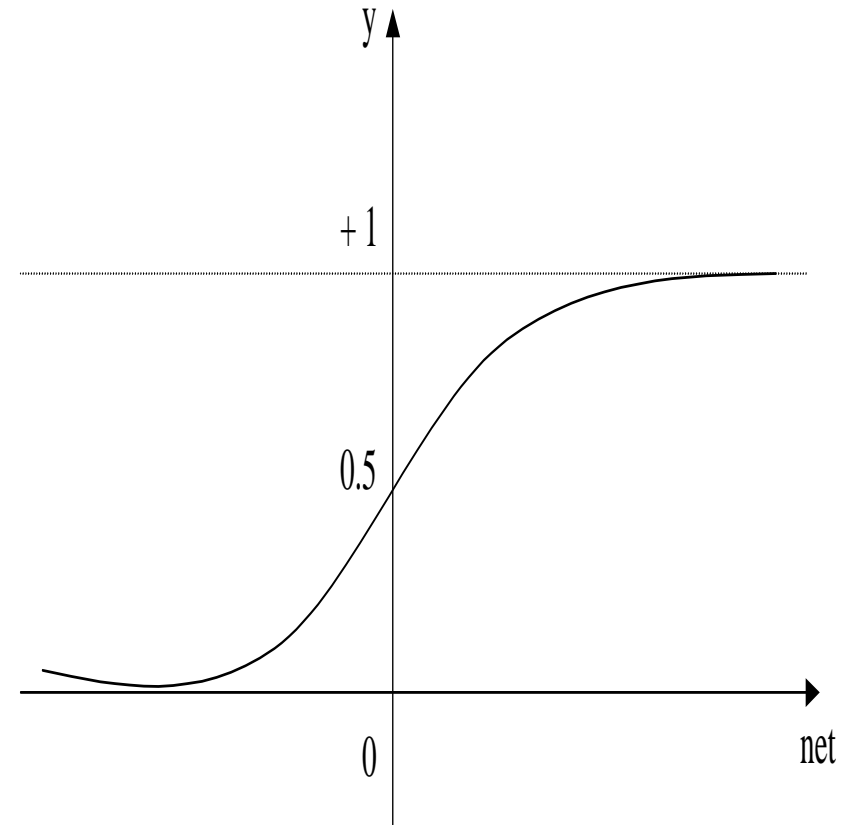
**Sigmoid Function II or  
Hyperbolic tangent  
Function**

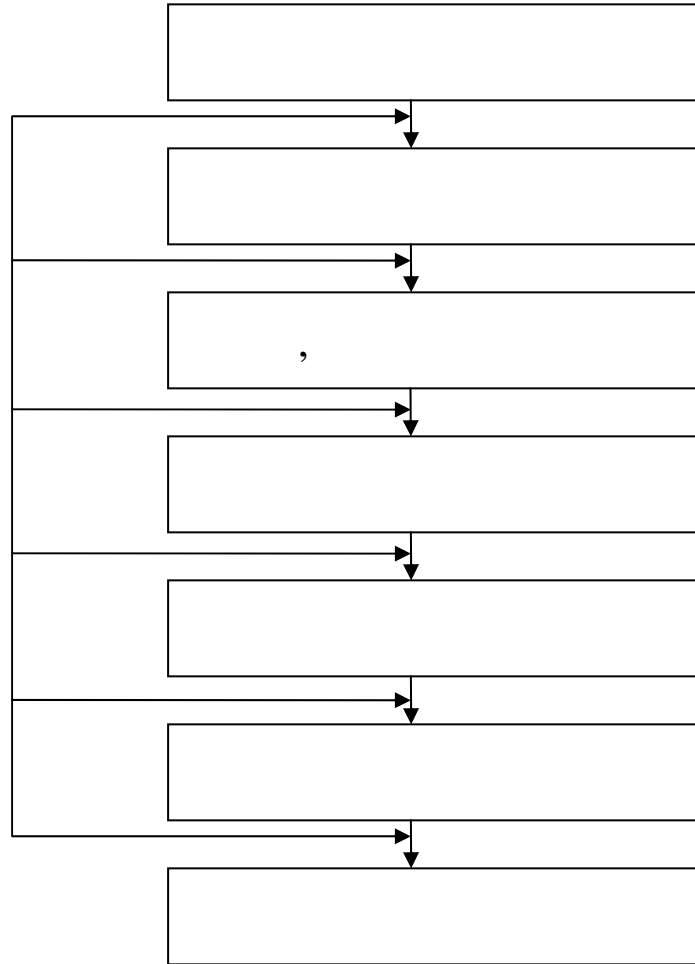
# (Transfer function)

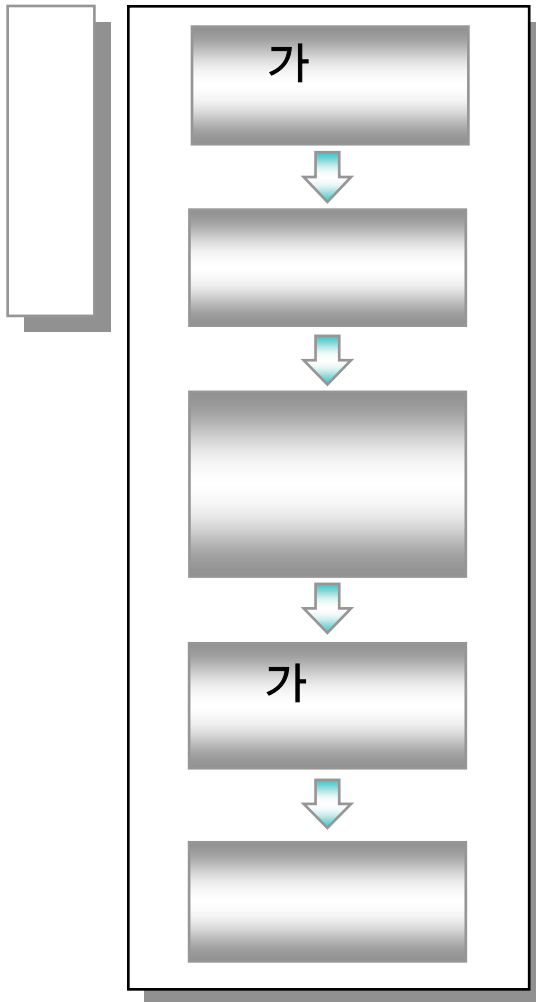
( )

## ■ Sigmoid Function

- 0 1
- 가 .
- S 가 .
- 







Learning: C:\WNSHELL2\EXAMPLES\WELECTRIC

File Run Options Help

**Training Graphics**

Training Set Average Error Epochs Elapsed  
 Test Set Average Error Intervals Elapsed  
 Error Factor Ranges Training Set Patterns  
 Error Factor Ranges Test Set Patterns

There are 365 training patterns.

learning events: 200750  
 learning epochs: 550  
 last average error: 0.0032530  
 minimum average error: 0.0020646  
 epochs since min. avg. error: 121

Automatically Save Training or

best training set     best test set

Check boxes above to display selected statistics. Training is

There are 0 test patterns.

**Training Set Average Error**

File Help

Refresh

Error

Epochs Elapsed

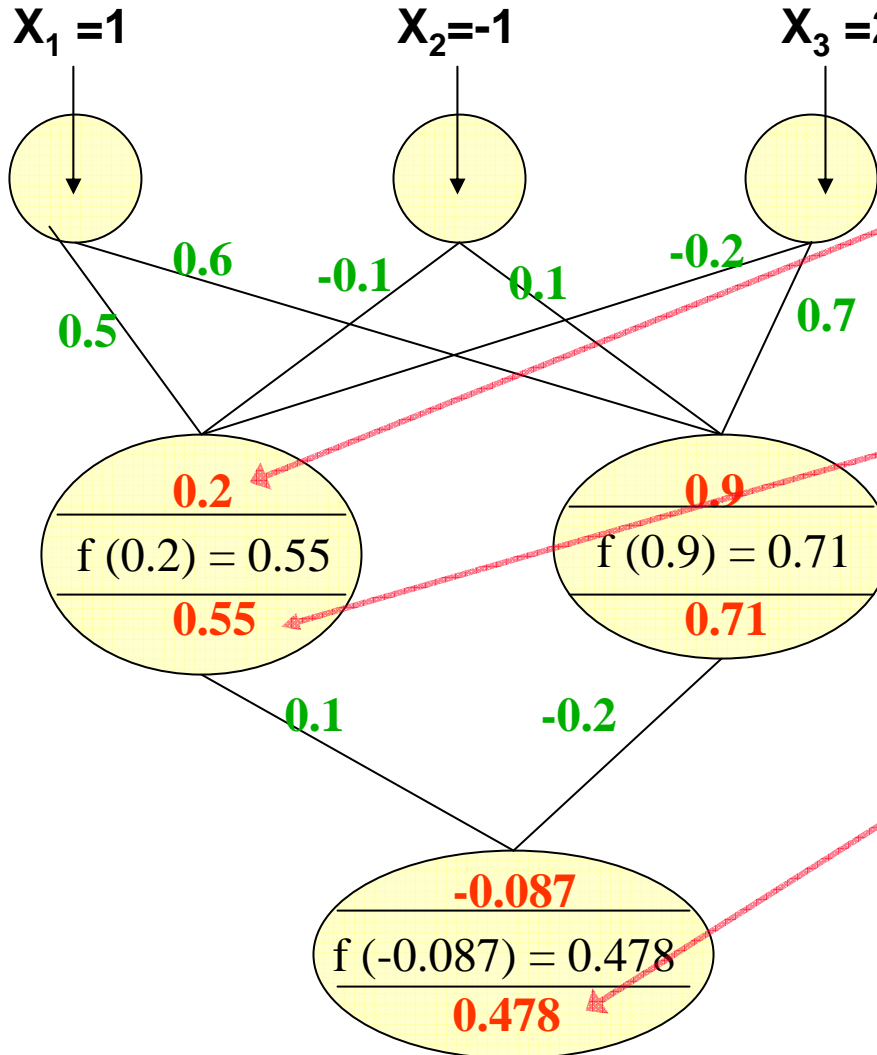


( )

Input:  $X_1 X_2 X_3$

Output:  $Y$

Model:  $Y = f(X_1 X_2 X_3)$



$$0.2 = 0.5 * 1 - 0.1 * (-1) - 0.2 * 2$$

$$f(x) = \frac{e^x}{1 + e^x}$$
$$f(0.2) = \frac{e^{0.2}}{1 + e^{0.2}} = 0.55$$

Predicted  $Y = 0.478$

Suppose Actual  $Y = 2$   
Then  
Prediction Error =  $(2 - 0.478) = 1.522$



.



● Rule of thumb

. ( .)

$\frac{1}{2}$  (2 +1)



=

(

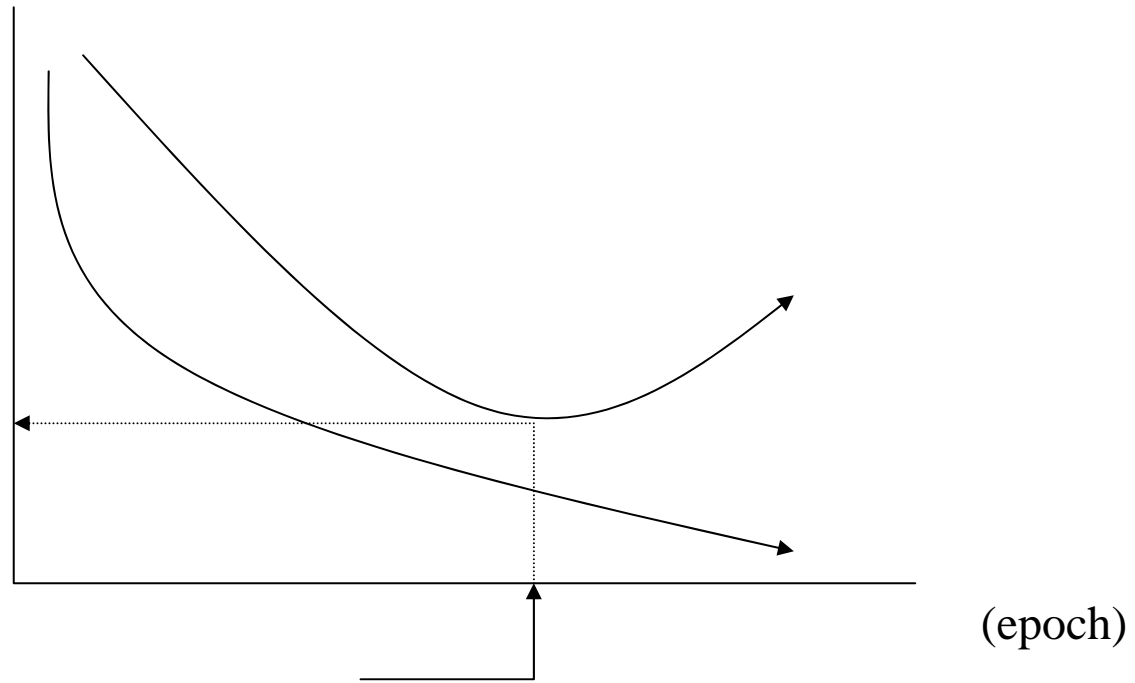
+

. ( ))



(Overfitting)

.





---

# **(Decision Tree) / (Inductive Learning or Rule Induction)**



**(decision rule)**

**(prediction)**





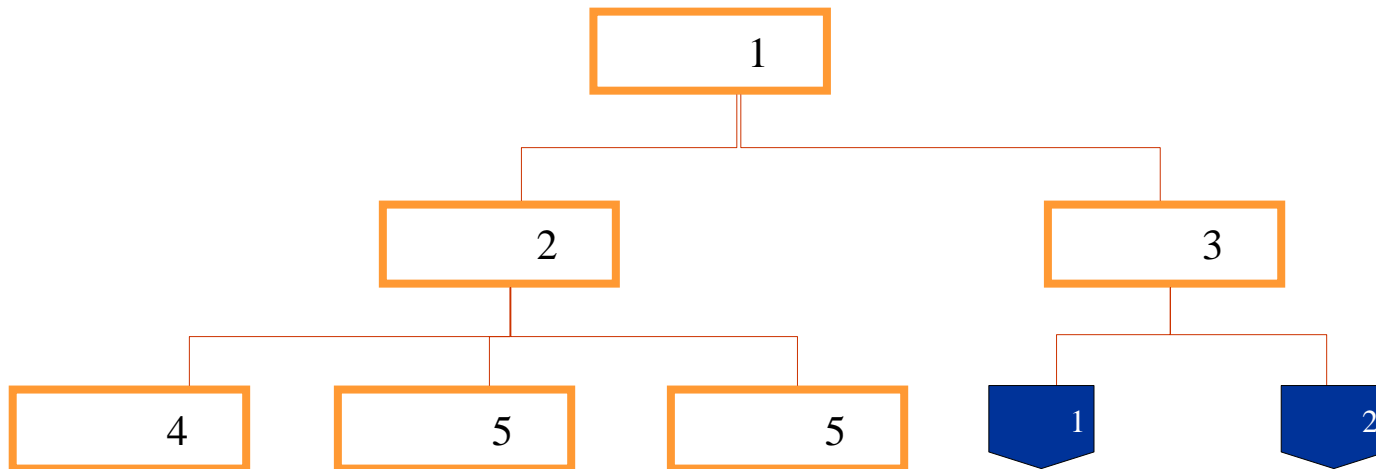
**(classification)**

.

**(Class)** ,  
**(Attribute)** .

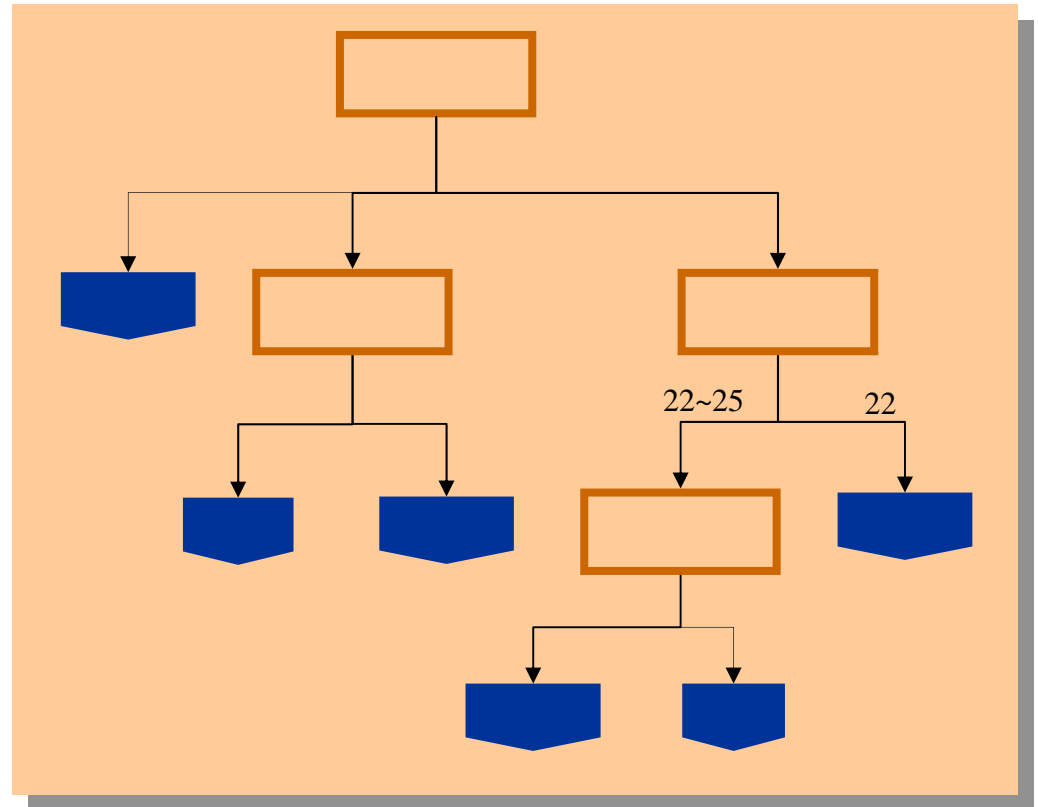
( )

- 
- (node)  가 (arc) —
- (leaf node)  가
- 가
- 가
- 



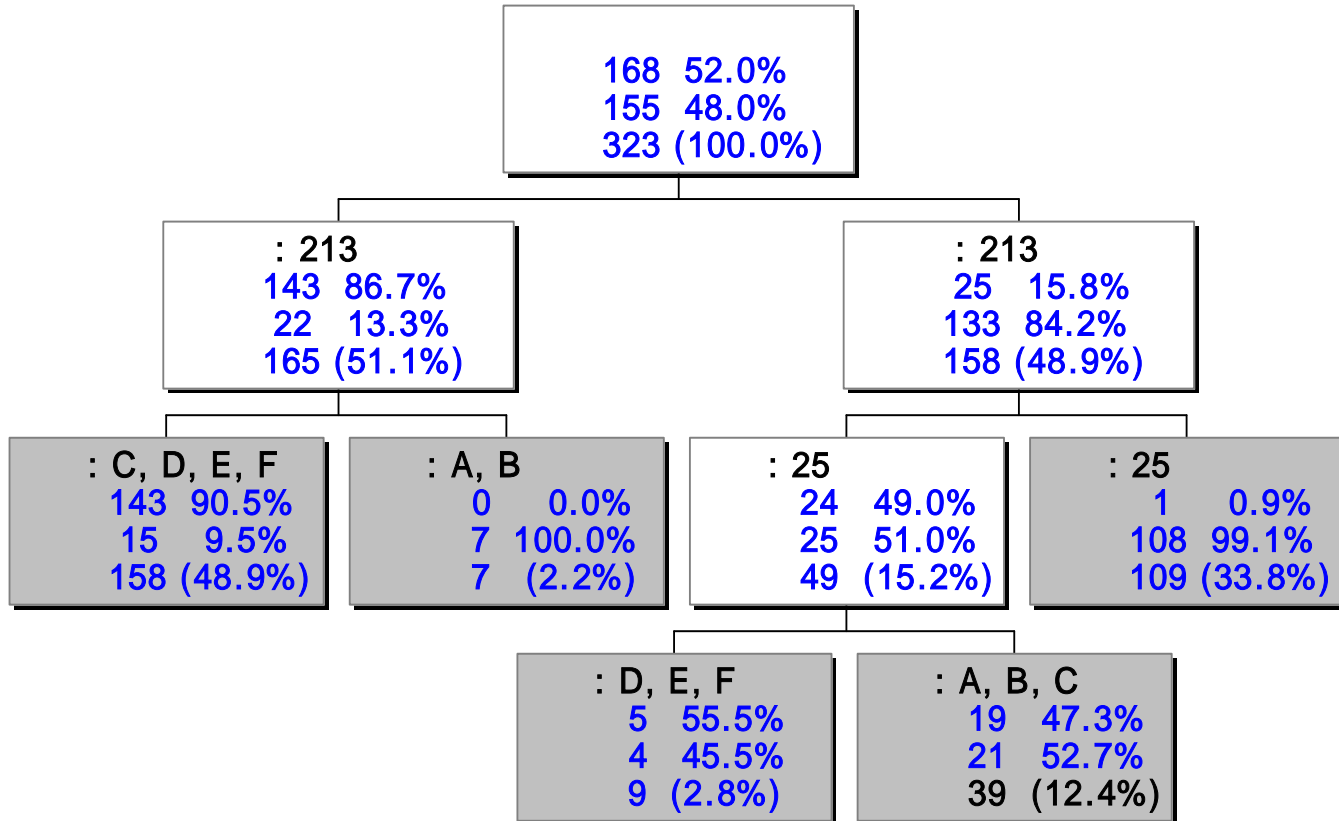
가?

- - : 22 ,  
22 ~ 25 ,  
25
  - : ,
  - : ,
  - : , , ,
- 가 , 가
- 가 , 가 22

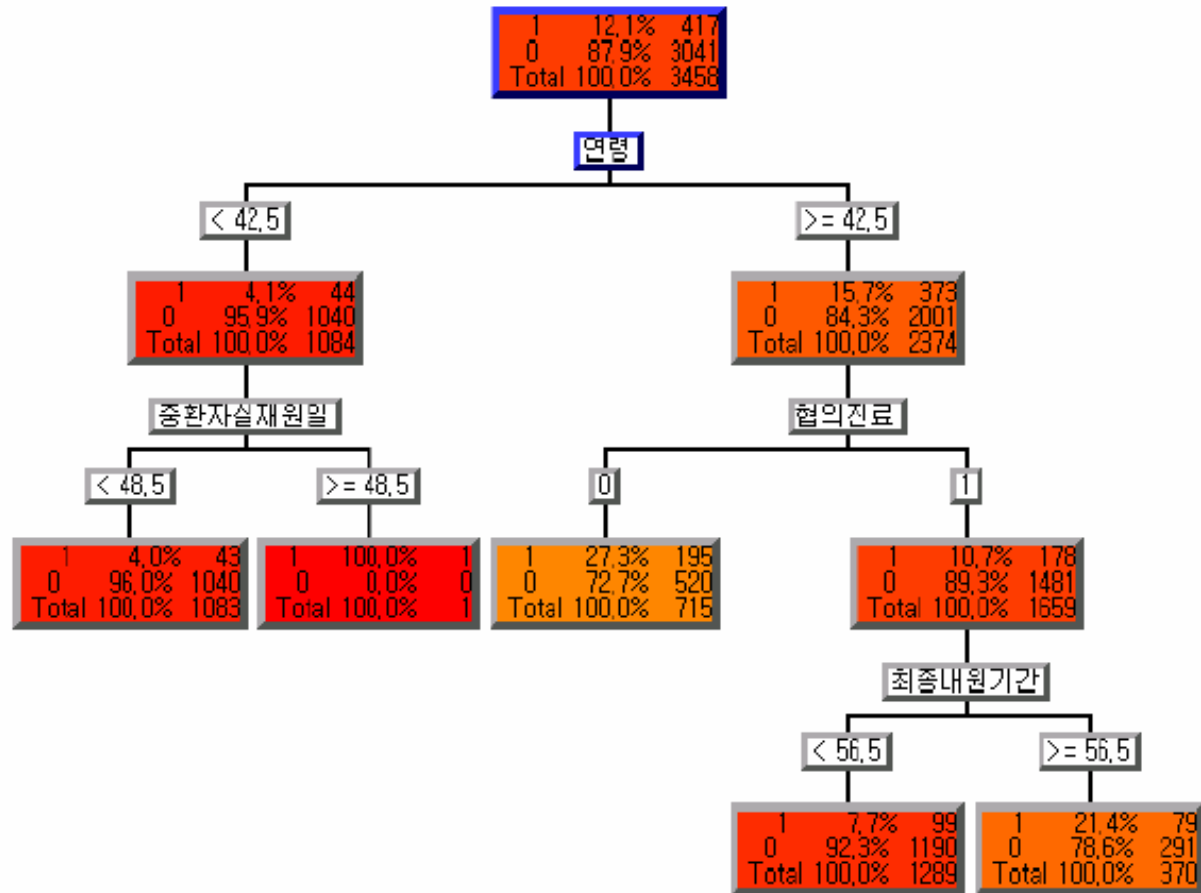




( )



( )



1:

2:

(

)

	CHAID	CART	QUEST
	' , '	' , '	' , '
	' , ' ( )	' , '	' , '
	- F		- F (Levene )
	(multiway)	(binary)	(binary)
가			

( )

	<b>C5.0 (Quinlan, 1996)</b>	<b>C&amp;R Tree (CART) (Breiman et al., 1984)</b>
	<b>(multiple)</b>	<b>2 (binary)</b>
	<b>(entropy)</b>	<b>( )</b>
가	가 ( )	가 ( )
가	-	
가 / <b>(minrec)</b>	<b>2 + 가 ≥ minrec</b>	<b>2 가 ≥ minrec</b>

( )

---

■ (Chi-Square statistic) p

- P- 가

■ (Gini index)

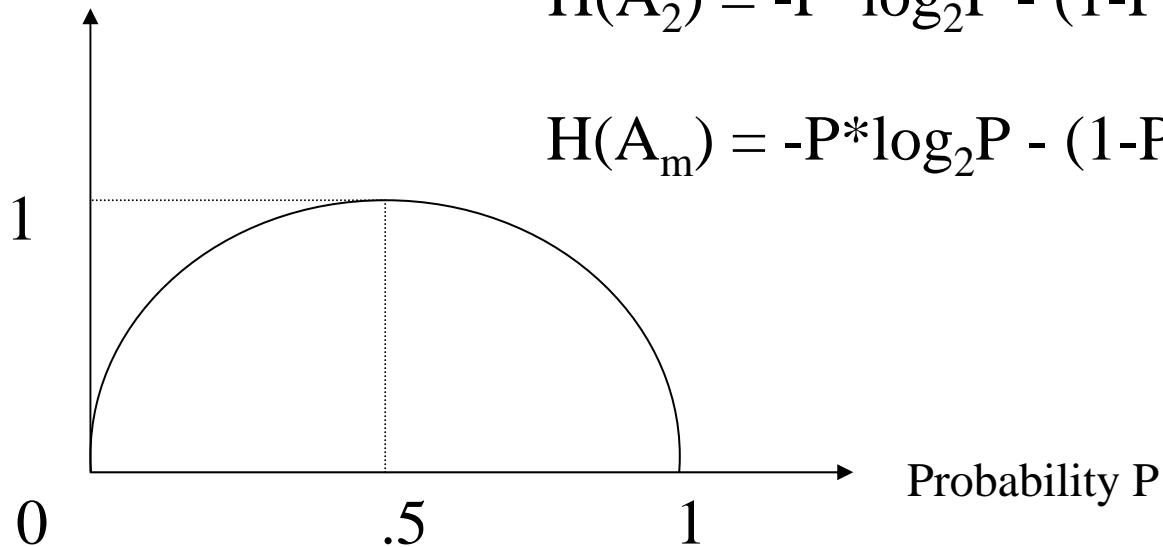
- 가

■ (Entropy index)

- . 가 가

# (Entropy)

Entropy in bits



$$H(A_1) = -P \cdot \log_2 P - (1-P) \cdot \log_2 (1-P)$$

$$H(A_2) = -P \cdot \log_2 P - (1-P) \cdot \log_2 (1-P)$$

$$H(A_m) = -P \cdot \log_2 P - (1-P) \cdot \log_2 (1-P)$$

$P = A$

,  $(1-P) = B$

---

- (Growing the Tree)

- (Splitting Criteria),
- (Stopping Rule)

- 가 (Pruning)

- 가 , 가 가

- 가(Validation)

- (Gain), (Risk), (Cost)

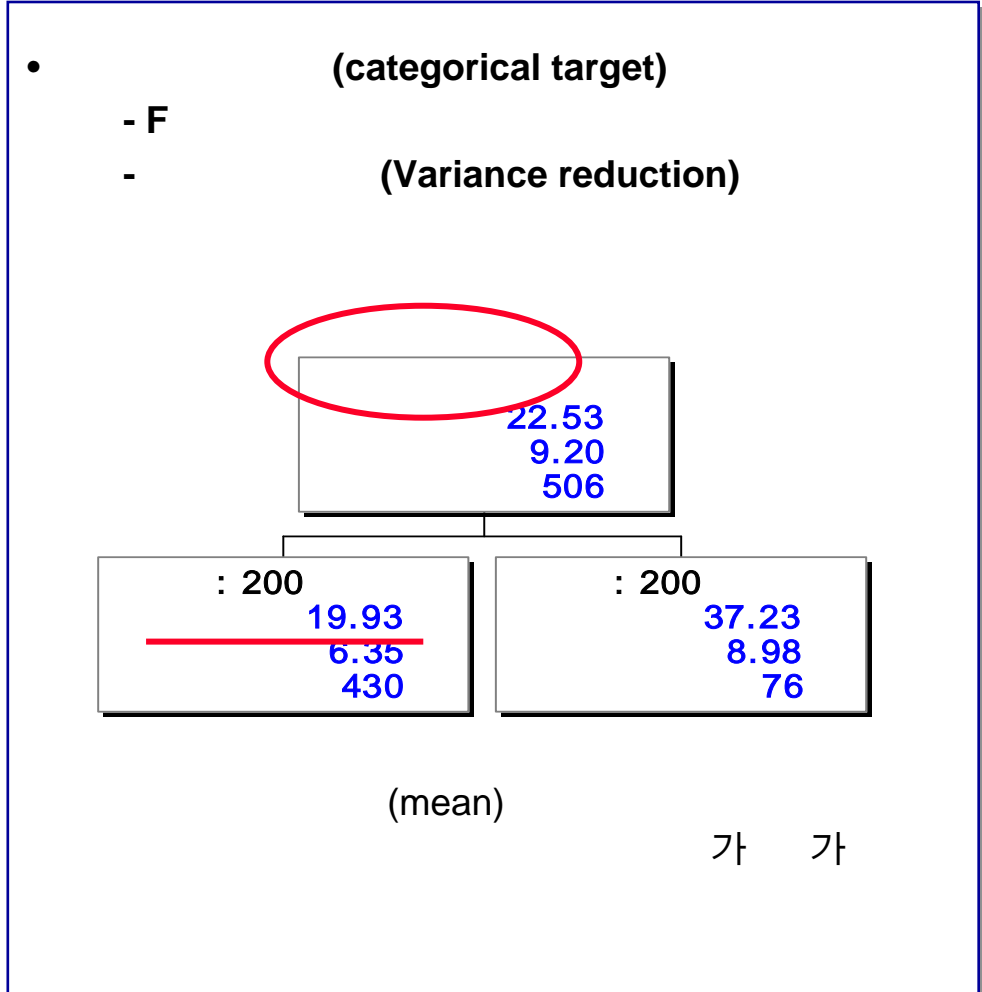
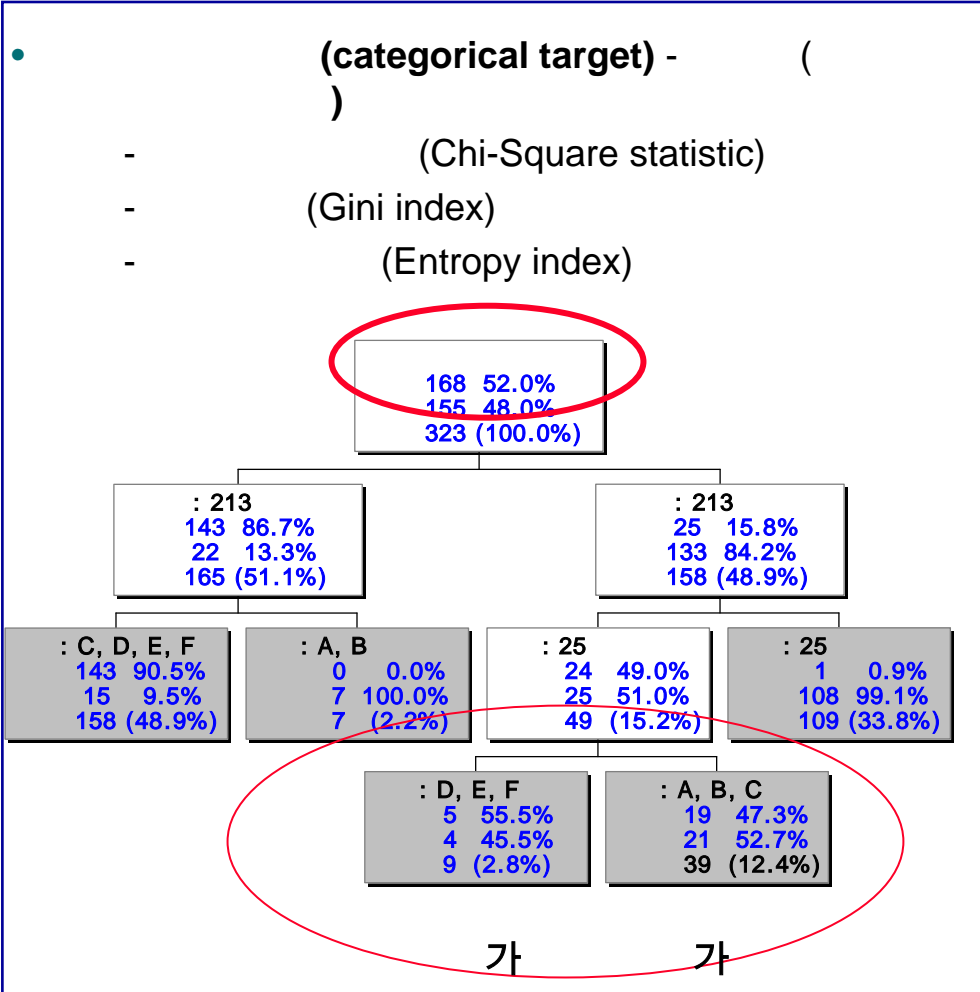
- (Classification, Prediction)

-

# (Splitting Criteria)

- (target variable)가 (interval) (categorical variable)

- (regression tree) , (classification tree) 가





,



•

,

,

가

.

•

.



•

가

가

.

---

- (Segmentation)

- 

( : , )

- (Classification)

- 

- 

- 

- 

- (Interaction)

- 

-

---

Neural Networks: SONN)  
(Kohonen Networks)

(Self-organizing

---

- ,

- 

(cluster)

- 가 가

- ,

, .

,



•

)

가 ( , ,  
가  
가



•

가 .



● 가



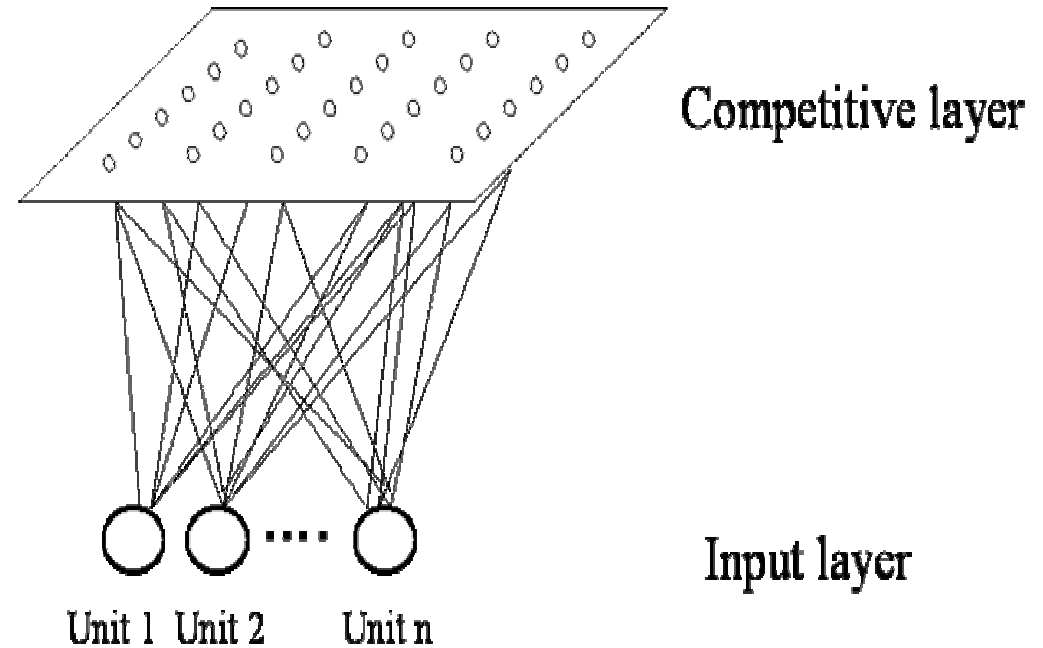
(k- , SONN)



# (Self-organizing Neural Networks: SONN)

- (Self-organizing Neural Networks: SONN) (Kohonen)

- (winner take all)



< > (SONN)  
(Kohonen Network)

# SONN

---

- - 
  - 
  - 
  - 
  - ( )
  -
- (feedforward flow)

# (Competitive Learning)

- input vector:

$$x = [x_1, x_2, x_3, \dots, x_m]^T$$

- weight vector:

$$w_i = [w_{i1}, w_{i2}, w_{i3}, \dots, w_{im}]^T, \quad i = 1, 2, \dots, l$$

- 가 가 가

(winner take all)

- best matching node

$$c = \arg \min_i \| x - w_i \|, \quad i = 1, 2, \dots, l$$

$$\text{same as } \| x - w_c \| = \min_i \{ \| x - w_i \| \}$$



# (Competitive Learning) ( )



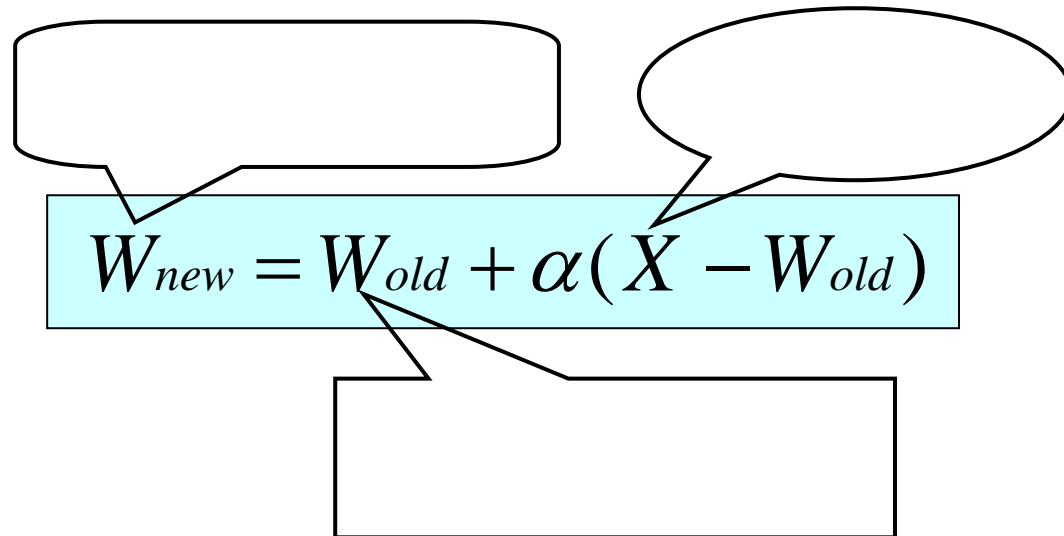
가 가 가

(Winner take all)

- 가가 가 ( )

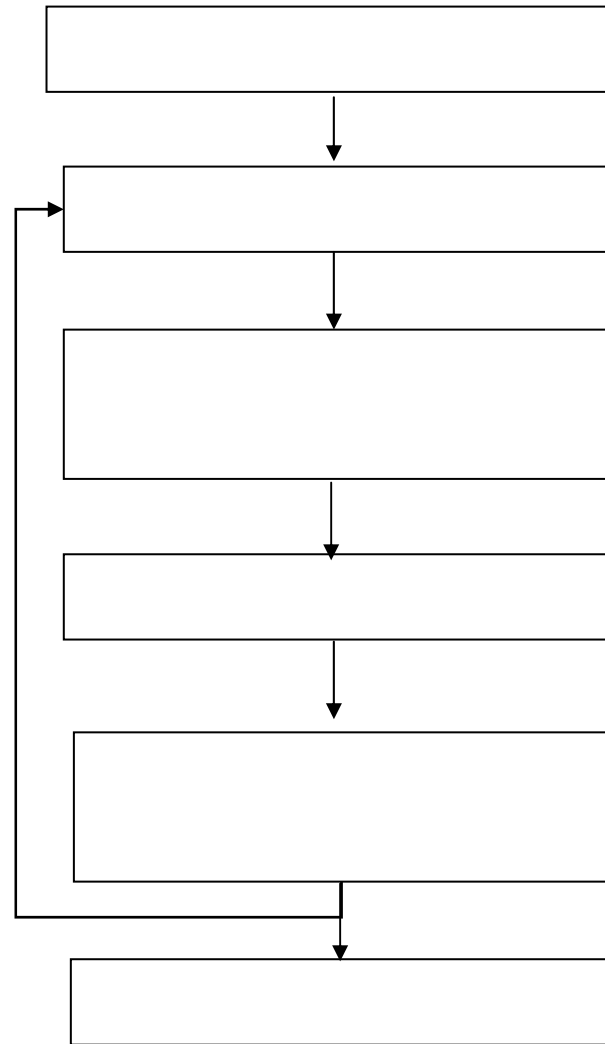
-

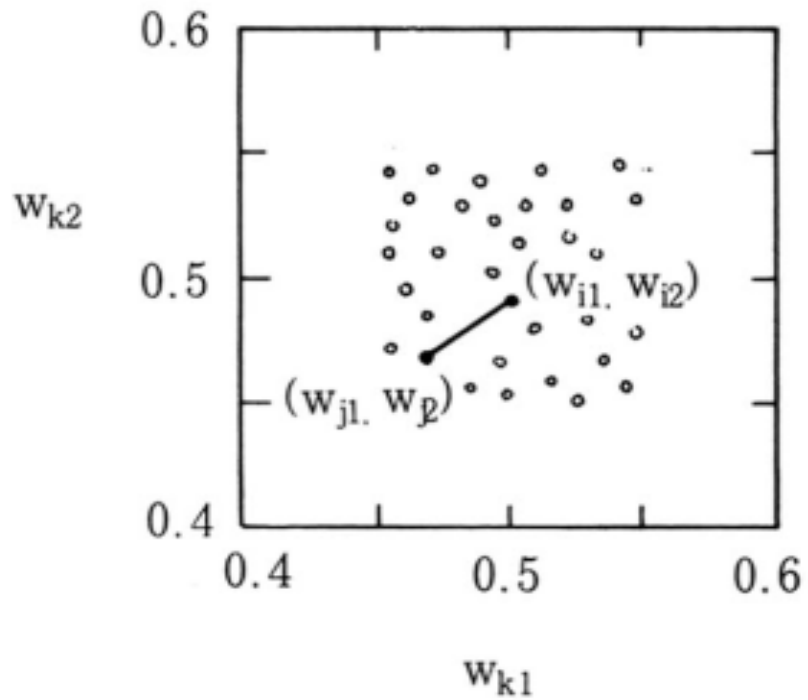
-



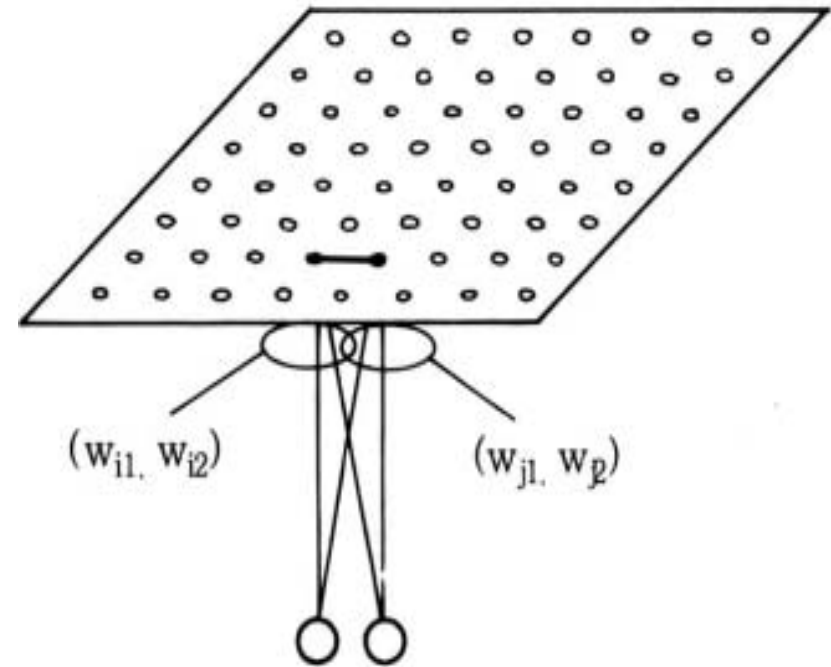
# SONN

---

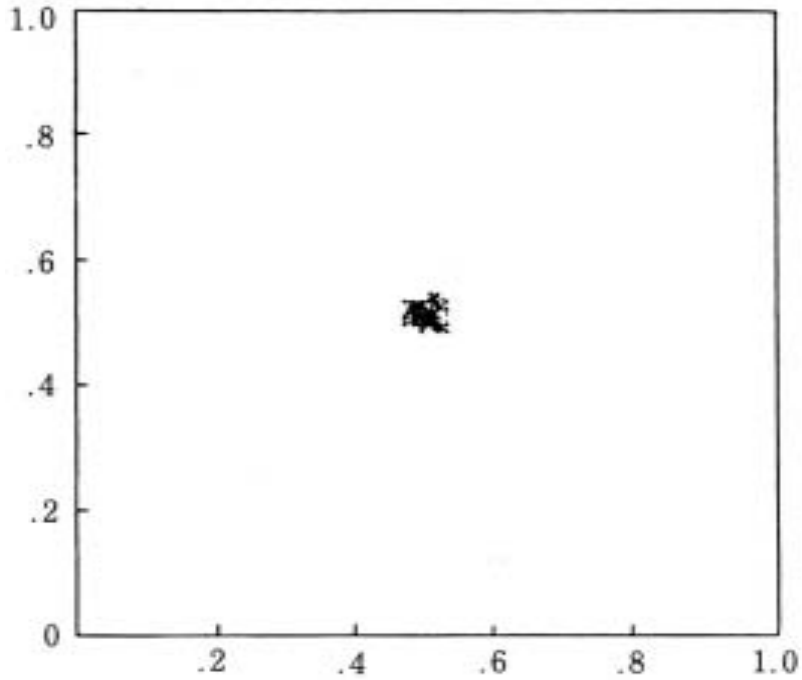




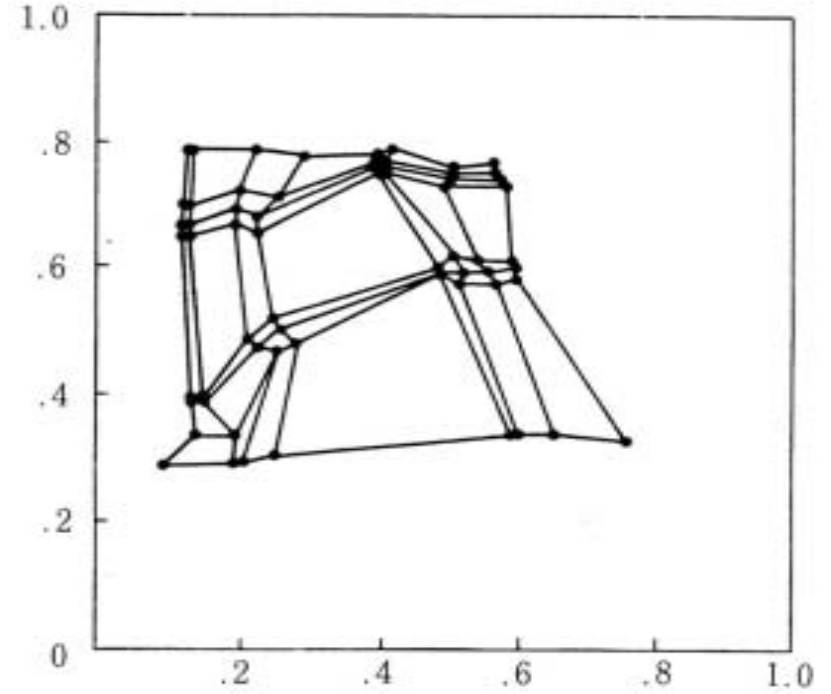
(a)



(b)

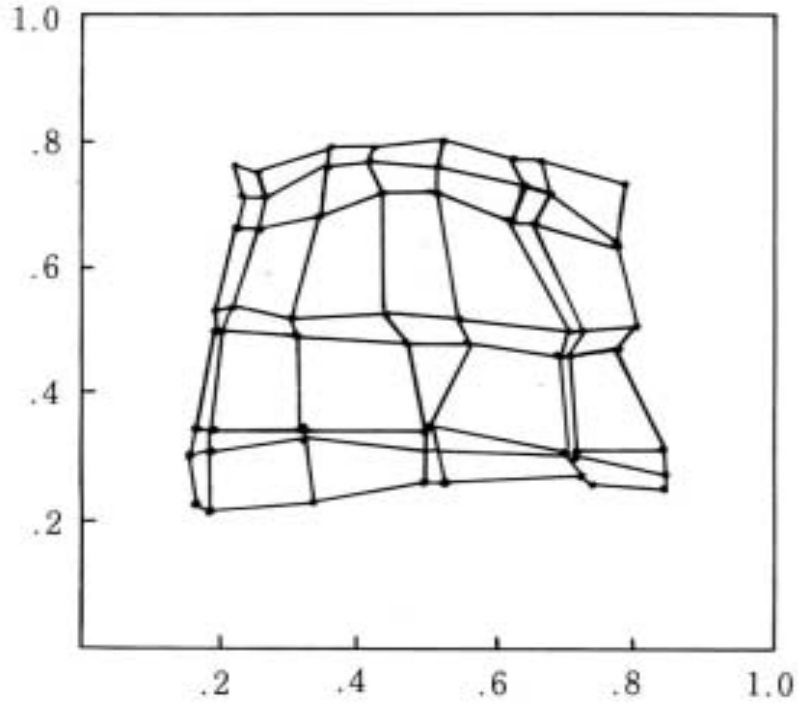


(a)



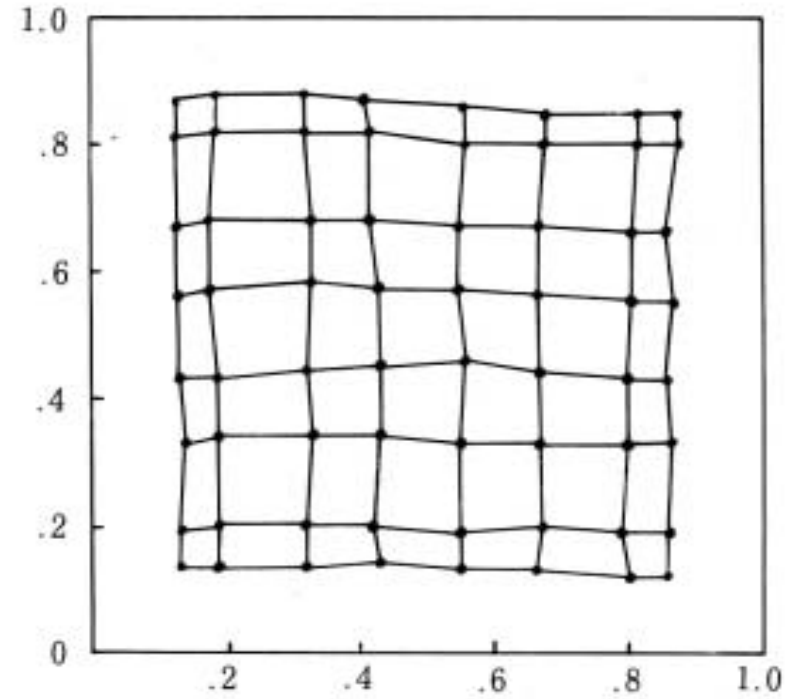
(b)

1000



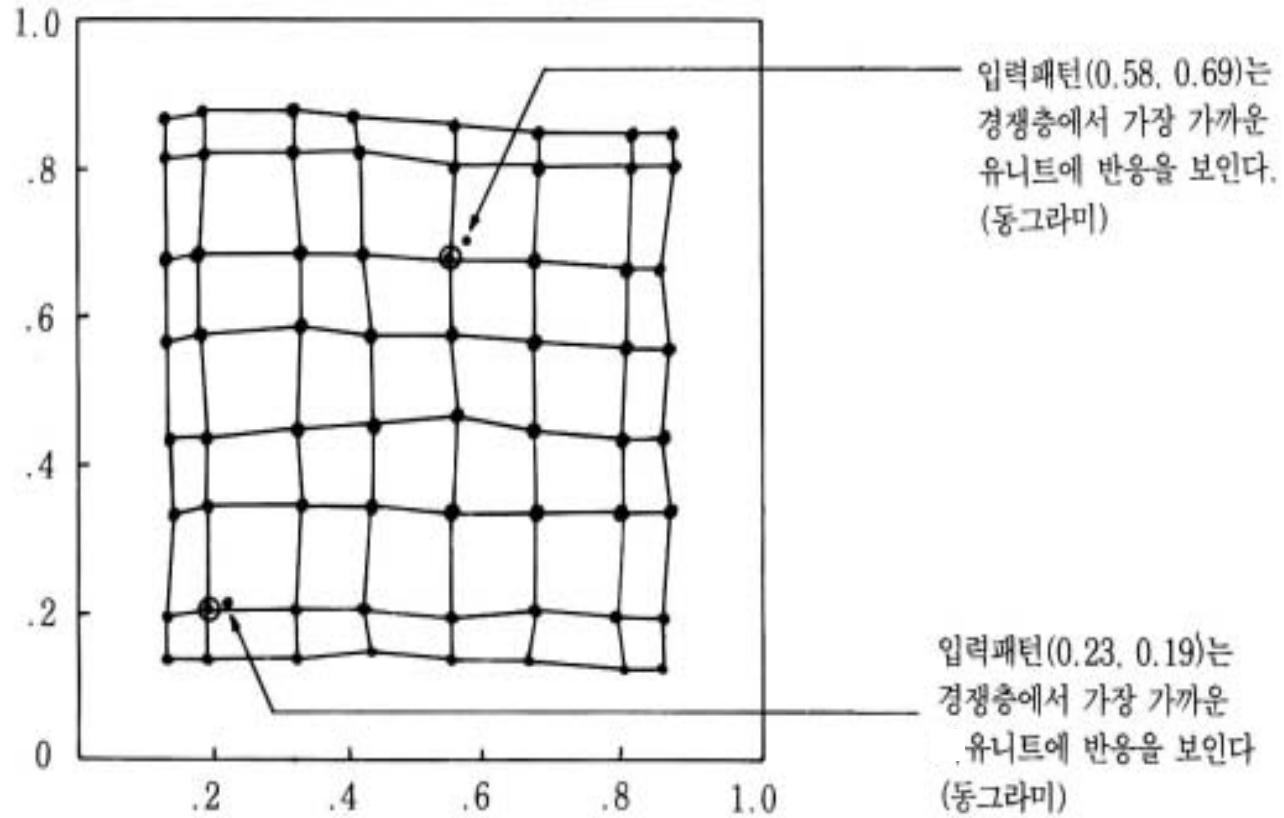
(a)

6000



(b)

20000



# SONN

---

- ( / )
- :
- , /

---

# (Association Rule)



# (Association Rule)

---

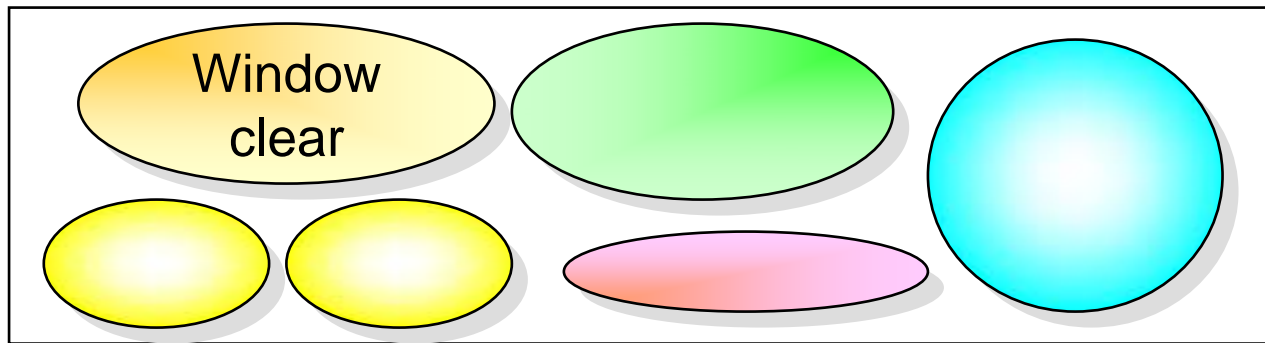
- Agrawal et al. 1993  
A B A=>B A가 B
- (affinity grouping)

# (Association Rule) ( )

?



EX) Products in Shop Cart (One trip, Together)



- 1) 가 가? 가?
- 2) 가? 가?
- 3) 가? 가?
- 4) 가? 가?

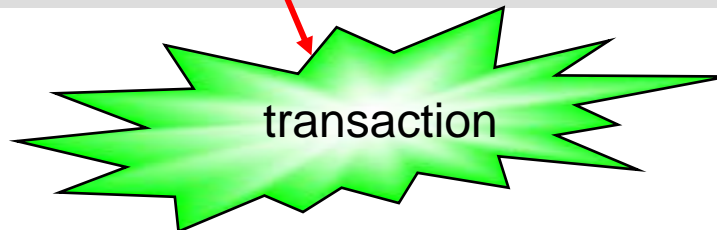
# (Association Rule) ( )

## ■ (transaction) (item)

- Market Basket Analysis product
- service offering .

Ex) Point-Of-Sale Transaction

customer	Set of products
1	,
2	,
3	, Window Cleaner

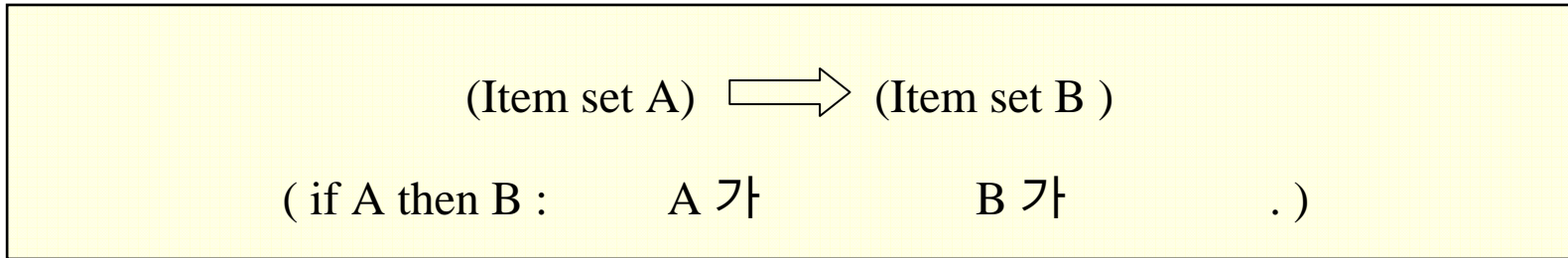


# (Association Rule) ( )

- (Association Rule)

- Item                    가                    Item

.



-

-

EX) 가

- 
- ( Cross Selling )
  - ( Inventory Display )
  - Catalog Design

- , , , ,

### List

ID	
1	, ,
2	, ,
3	,
4	,
5	, , ,
6	

### Co-occurrence of Product( )

	4	2	3	1	2
	2	2	3	0	1
	3	3	4	0	2
	1	0	0	2	0
	2	2	2	0	2

( )

1.

- , 가 combination
- , , .

2.

- .

?

,

.

---

■ (Support)

- X Y 가  
가 ?

$$S = P(X \cap Y) = \frac{X \quad Y}{(N)}$$

-



( )

■ (Confidence)

- X Y가  
가 ?

$$C = P(Y | X) = \frac{P(X \cap Y)}{P(X)}$$
$$= \frac{X \quad Y}{X}$$

- 
-

( )

■ ( Lift / improvement )

- X 가 Y Y가  
?

$$L = \frac{P(Y | X)}{P(Y)} = \frac{P(X \cap Y)}{P(X)P(Y)} = \frac{X \ Y}{X \times Y}$$

Lift		
1		
> 1		
< 1		,

( )

■

1. 가  
·  
- 가 ,
2. 가  
·
3. Lift 가 .  
가  $X \Rightarrow Y$  Y  
가 .  
- Lift 1 가 .

( )

List

ID	
1	, ,
2	, ,
3	,
4	,
5	, , ,
6	



가 50%

50%	Transaction	
=>	1,2,5	75 %
=>	1,4,5	75 %
=>	1,4,5	100 %

$Lift = P( \quad | \quad ) / P( \quad ) = 1 / (4/6) = 1.5$

\* \_\_\_\_\_ :

(100%)

- 가

(50%)

- 가

가

1.5

( )

- (URL)
  - “ ” 45%
- “ ” 4 5 , 9 10
- “A B가 가 C 가 .”

,

■

1. 가 .
2. .

■

1. 가 .
2. 가 .
3. DBMS
4. 가 가 . 가 .