



# **An Evolutionary Approach to Credit Scoring**

---

**Prof. Shinn-Ying Ho**

**Department of Biological Science and Technology  
and Institute of Bioinformatics  
National Chiao Tung University**

**e-mail: [syho@mail.nctu.edu.tw](mailto:syho@mail.nctu.edu.tw)**

**Tel.:886-3-5712121-56905**



# Outline

- Introduction
- Problem Definition and Related Work
- Design of Accurate Credit Scoring Model
  - Intelligent Genetic Algorithm (IGA)
  - The Evolutionary Design Approach
- Experimental Result
- Discussions
- Conclusions



# Introduction

- **Basel Capital Accord**
  - In 1988, the Basel Committee on Banking Supervision adopted a capital adequacy framework for internationally active commercial banks. The Accord specifically created capital requirements for the credit risk in banking assets.
  
- **Basel 2 Accord**
  - The framework includes three pillars:
    - Minimum capital requirements
    - Supervisory review
    - Market discipline



## Introduction

- The new accord contains several revisions that should better align capital requirements with actual bank risk included market and **credit risk**.
- In order to fulfill the requirements of complex risk analysis, the computational intelligence plays an important role for developing advanced risk management system.



# Introduction

- Importance of Credit Scoring
  - Between 1985 and 1996, the number of personal bankruptcy cases filed in US rose from 341,000 to 1.1 million. The rate of bankruptcy per 1000,000 adults increases from 203 to 596 (Nelson, 1999).
  - However, the competition in the consumer credit market has become intense.
  - Minimizing the credit risk can increase the total gain.



# Introduction

- What is Credit Scoring
  - Evaluate the credit worthiness of customers and predict the possibility of default loans.
- Applications
  - Financial institutions
  - Insurance institutions
  - Sell markets



# Introduction

- Credit Scoring vs. Credit Rating
  - Both of them measure the credit worthiness of customers.
  - Credit Scoring
    - Predict the worthiness, and then accept the loan application for “good credit” or reject it for “bad credit”.
  - Credit Rating
    - Classify the worthiness into levels such as “AAA”, “AA”, and “A”, and assign different loan quota for each level.



# Introduction

- Advantages of Machine Learning Approaches
  - Eliminate the risk from human defect.
  - Build customize models according to the constitutions of banks.
  - The banks can rebuild the model themselves when the environment changed.





# Introduction

- Most of the machine learning approaches for credit scoring aim on the classification accuracy but lack of interpretation of learned models.
- In this study, we proposed an evolutionary approach to designing a fuzzy classifier for solving credit scoring problems.
- The proposed fuzzy rule-based classifier can provide both high explanation ability and high accuracy.
- All design parameters of the fuzzy classifier are optimized using an inheritance intelligent genetic algorithm.



# Problem Definition

- Goal
  - Establish an accurate scoring model according to personal information and historical records of existing customers.
  - Use it to predict customers' credit worthiness.
- The characteristics of credit scoring
  - Two classes: “good credit” and “bad credit”
  - The number of “good credit” samples may be much more than that of “bad credit” ones.
  - Different misclassification cost for the two classes.

# Problem Definition

- Type 1 error: classify “bad credit” samples as “good credit”
- Type 2 error: classify “good credit” samples as “bad credit”
- Type 1 error is more critical than type 2 error.
- The cost matrix

Cost		Predict class $\hat{y}$	
		Good	Bad
Predict class $y$	Good	0	1
	Bad	5	0

## Problem Definition

- Let  $X = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\}$  be the set of  $m$  training samples. Each  $\mathbf{x}_i \in \mathcal{R}^n$  represents an  $n$ -D feature vector, and  $y_i \in \{0, 1\}$  is its real class (0:good, 1:bad).
- The objective is to design a decision function  $\hat{y} = f(\mathbf{x})$ ,  $\hat{y} \in \{0, 1\}$  that minimize the total cost:

$$\min \sum_{i=1}^m \text{Cost}(y_i, \hat{y}_i)$$



# Problem Definition

- Scoring functions
  - Design of a function  $s(\mathbf{x})$  that maximize the between-group variance relative to the with-in group variance.
  - “Good credit” samples have higher scores than “bad credit” ones.
  - The decision function  $f(\mathbf{x})$  is defined as

$$f(\mathbf{x}) = \begin{cases} 1, & \text{if } s(\mathbf{x}) < Z \\ 0, & \text{if } s(\mathbf{x}) \geq Z \end{cases}$$

where  $Z$  is the threshold minimized the cost.



## Related Work

- Human Experience
  - Scorecard
- Statistic Method
  - Logistic Regression
  - Discriminant Analysis
- Non-parametric Models
  - *K*-Nearest Neighbor Rule
  - Decision Tree
  - Neural Network



# Comparison of Models

Models	Accuracy	Interpretable	Quantitative Score
Scorecard	Poor	Yes	Yes
Statistic methods	Poor	No	Yes (DA)
<i>k</i> -NN	Good	No	No
Decision trees	Good	Yes	No
Neural networks	Good	No	Yes
Fuzzy rule-based classifier	Good	Yes	Yes



# Scorecard

項目	評分依據 (請勾選答案)
年齡	<input type="radio"/> 未滿20歲 <input type="radio"/> 20歲~25歲 <input type="radio"/> 26歲~30歲 <input type="radio"/> 31歲~35歲 <input type="radio"/> 36歲~45歲 <input type="radio"/> 46歲~55歲 <input type="radio"/> 56歲以上
年收入 (含配偶)	<input type="radio"/> 本人(含)配偶合計所得120萬元以上 <input type="radio"/> 本人(含)配偶合計所得100萬元以上 <input type="radio"/> 本人(含)配偶合計所得80萬元以上 <input type="radio"/> 本人(含)配偶合計所得60萬元以上 <input type="radio"/> 本人(含)配偶合計所得50萬元以上 <input type="radio"/> 本人(含)配偶合計所得30萬元以上 <input type="radio"/> 其他
職業	<input type="radio"/> 政府機關，軍警單位，公營事業，公私立學校等正式人員或醫師，律師，會計師，建築師。 <input type="radio"/> 上市/上櫃公司股票公開發行公司，或排名全國一千大之民營企業編制內之職員。 <input type="radio"/> 非前述民營事業人員，農、林、漁、牧事業及其他自由業。 <input type="radio"/> 其他
工作年資	<input type="radio"/> 已工作10年以上 <input type="radio"/> 工作6年以上~10年 <input type="radio"/> 工作3年以上~6年 <input type="radio"/> 工作1年以上~3年 <input type="radio"/> 未滿1年

住宅(房地產) 狀況	<input type="radio"/> 本人或配偶之房屋未設定抵押權 <input type="radio"/> 本人或配偶之房屋已設定抵押權給銀行 <input type="radio"/> 本人或配偶未擁有不動產
家庭狀況	<input type="radio"/> 已婚，有子女者 <input type="radio"/> 已婚，無子女者 <input type="radio"/> 未婚
金融機構 往來情形	<input type="radio"/> 銀行存款往來半年以上，且平均存款餘額達十五萬元以上。 <input type="radio"/> 銀行存款往來半年以上，且平均存款餘額達十萬元以上。 <input type="radio"/> 銀行存款往來半年以上，且平均存款餘額達五萬元以上。 <input type="radio"/> 銀行存款往來3個月以上，且平均存款餘額達五萬元以上。 <input type="radio"/> 銀行存款往來未滿三個月。
信用卡 往來情形	<input type="radio"/> 曾有逾期催繳之情形。 <input type="radio"/> 目前正常使用的信用卡張數1~2張。 <input type="radio"/> 目前正常使用的信用卡張數3~4張。 <input type="radio"/> 目前正常使用的信用卡張數達5張以上。
信用卡 卡齡	<input type="radio"/> 未滿一年。 <input type="radio"/> 持有信用卡達1~3年。 <input type="radio"/> 持有信用卡達3~5年。 <input type="radio"/> 持有信用卡達5年以上

(From <http://tw.money.yahoo.com/edu/credit.html>)





# Scorecard

## Score and Rating

分數級距	等級	(申貸金額)消費性貸款授信額度
95分(含)以上	您的信用頂尖! 為銀行樂於服務的白金尊榮客戶	申貸金額之每年攤還本息, 不超過年收入之50%均可承作
85分(含)以上	您的信用卓越! 為銀行樂於洽辦的超優質客戶	申貸金額之每年攤還本息, 不超過年收入之45%均可承作
75分(含)以上	你的信用良好! 為銀行樂於服務的優質客戶	申貸金額之每年攤還本息, 不超過年收入之40%均可承作
65分(含)以上	您的信用正常! 是銀行樂於服務的客戶	申貸金額之每年攤還本息, 不超過年收入之35%均可承作
55分(含)以上	您的信用尚待培養建立! 未來值得期待	申貸金額之每年攤還本息, 不超過年收入之25%均可承作

(From <http://tw.money.yahoo.com/edu/credit.html>)



# Scorecard

## ■ Drawbacks

- Static score card system cannot response in real time when the environment changed
- No “generic scorecard” is suitable for all banks
- The design depends on the scorecard providers. The cost is very high and the banks have less autonomy.

# Logistic Regression

- Let  $p=P(y=1)$  be the probability of “bad credit” and its linear regression model is defined as follows:

$$\ln \frac{p}{1-p} = g(x) = a_0 + a_1x_1 + \dots + a_nx_n$$

- If  $g(x)>0$  (i.e.  $p>0.5$ ), then  $y=1$ ; otherwise,  $y=0$ .
- $g(x)$  can be a linear or a non-linear model.

# ➤ Discriminant Analysis

- The discriminant analysis derives a linear combination of input variables

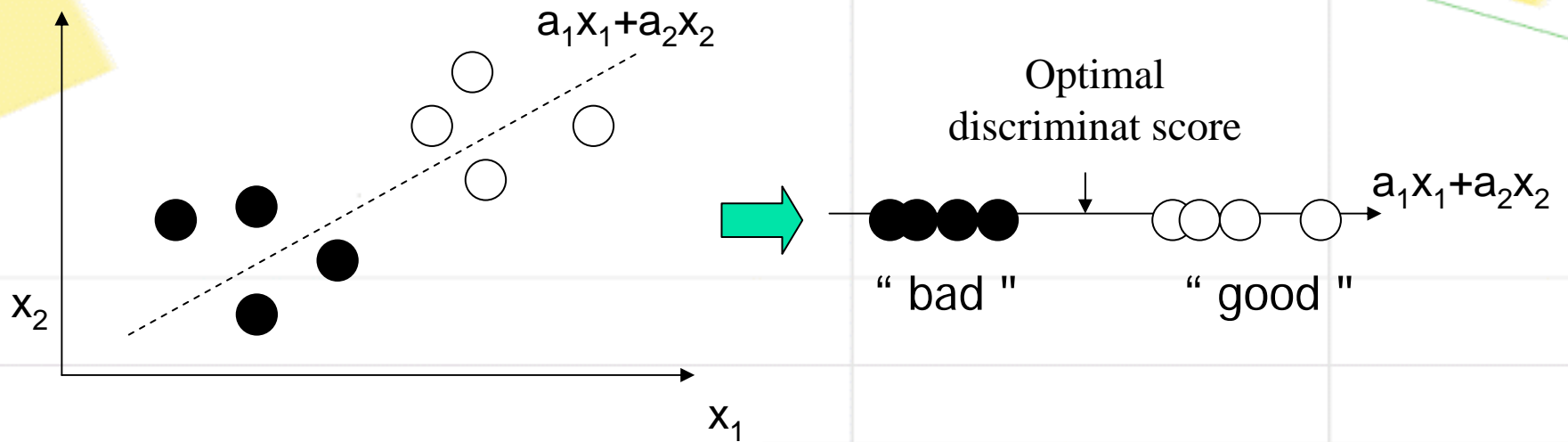
$$Z = a_1x_1 + a_2x_2 \dots + a_nx_n$$

where  $Z$  is the discriminant score.

- Maximize the between-group variance relative to the within-group variance.

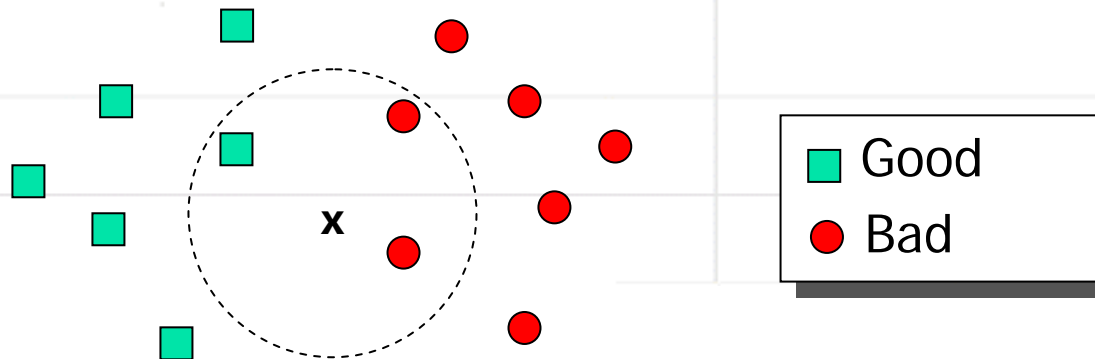


# Discriminant Analysis



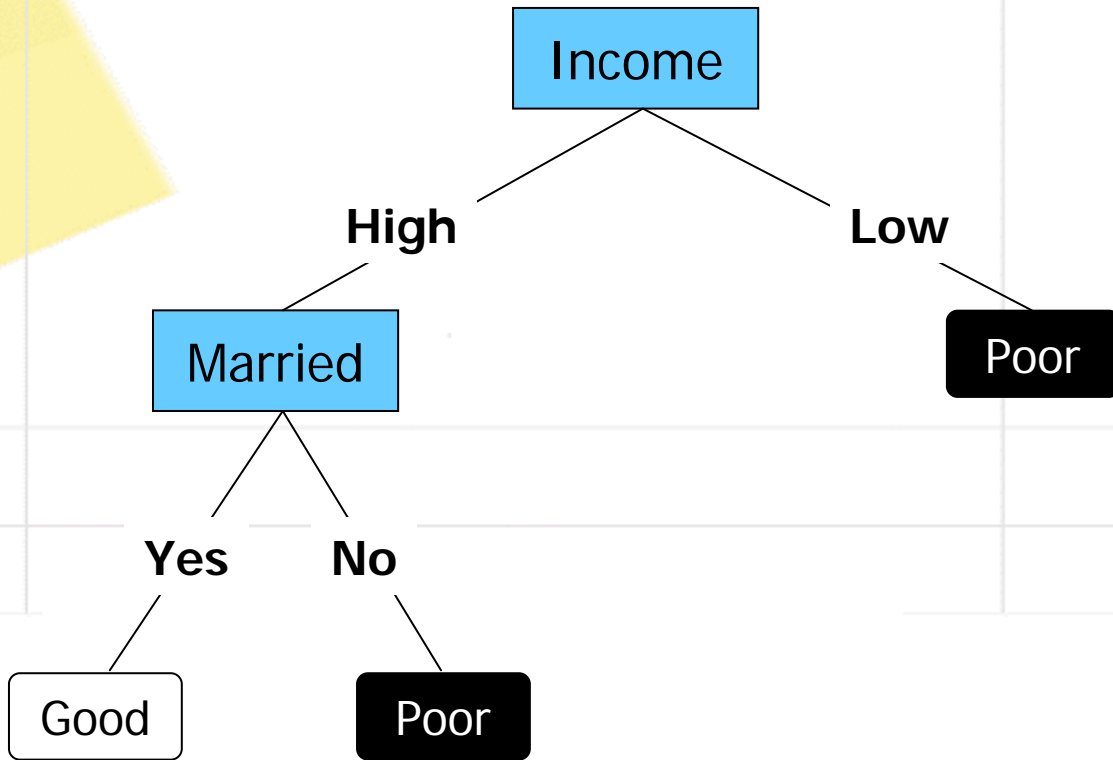
## ➤ *K*-Nearest Neighbor Rule

- Determine the class according to the voting result of the *k* nearest neighbors of given pattern.
- Example: pattern *x* is classified as “bad credit”.





# Decision Tree

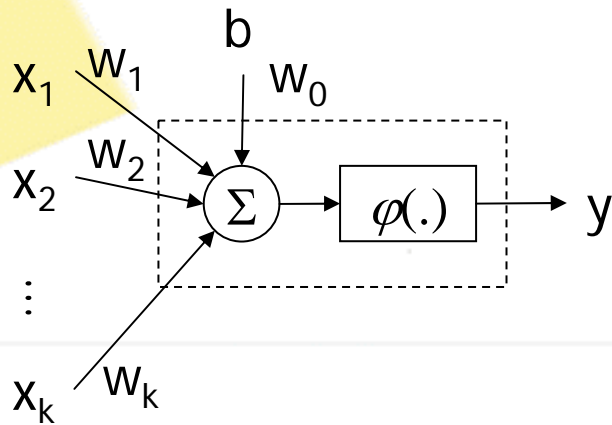


Name	Debt	Income	Married?	Risk
Joe	High	High	Yes	Good
Sue	Low	High	Yes	Good
John	Low	High	No	Poor
Mary	High	Low	Yes	Poor
Fred	Low	Low	Yes	Poor



# Neural Network

Neuron



$\Sigma$ : summation

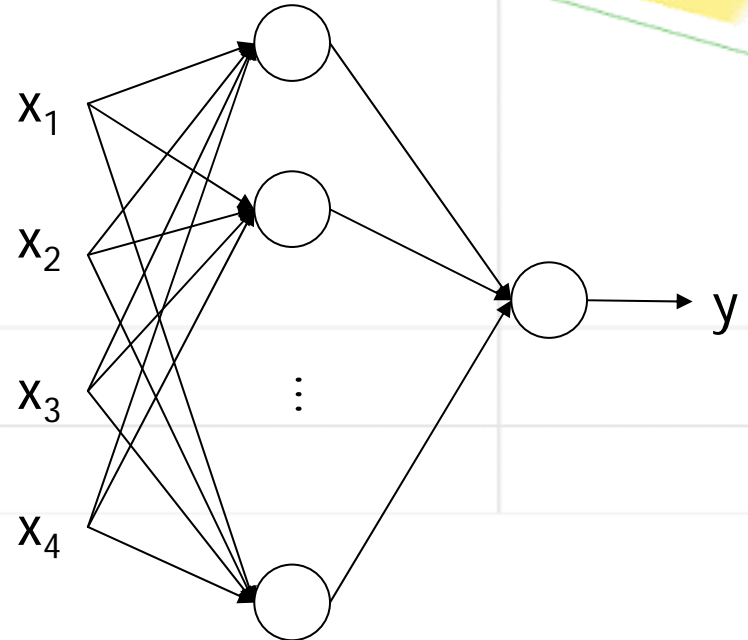
$b$ : bias

$\varphi(\cdot)$ : activation function

Input layer

Hidden layer

Output layer







## **Design of Accurate Credit Scoring Model**

- **Intelligent Genetic Algorithm**
- **The Evolutionary Design of Fuzzy Rule-Based Classifier**



## ➤ Design of Credit Scoring Model

---

- We proposed an evolutionary approach to solve credit scoring problem.
- All the design parameters of the fuzzy classifier are optimized using an inherent intelligent genetic algorithm.

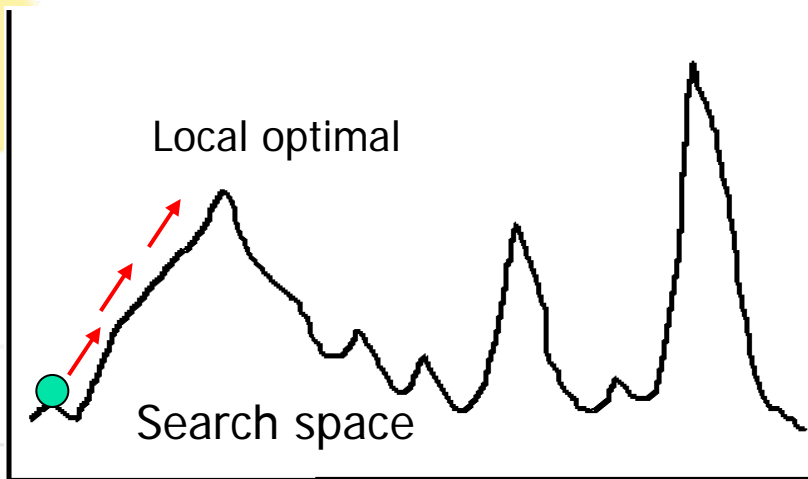
# Intelligent Genetic Algorithm

## ➤ Genetic Algorithm (GA)

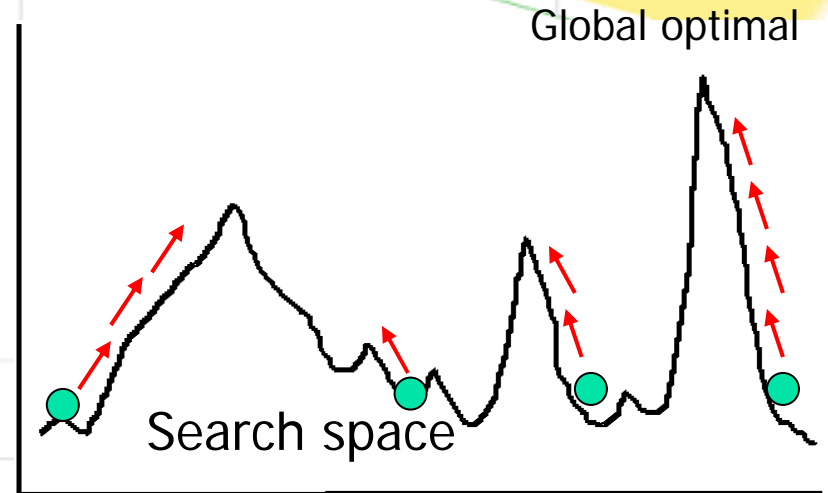
- Advantages
  - Robust
  - The optimization mechanism is independent of the problem.
  - Global search
    - It deals with the balance between exploration and exploitation in a search space.
  - Flexible
    - GA can easily integrate with heuristics to solve problem more efficiently.

# Intelligent Genetic Algorithm

## GA vs. Numerical Method



Numerical method



Genetic Algorithm

# Intelligent Genetic Algorithm

## ➤ Flowchart of GA

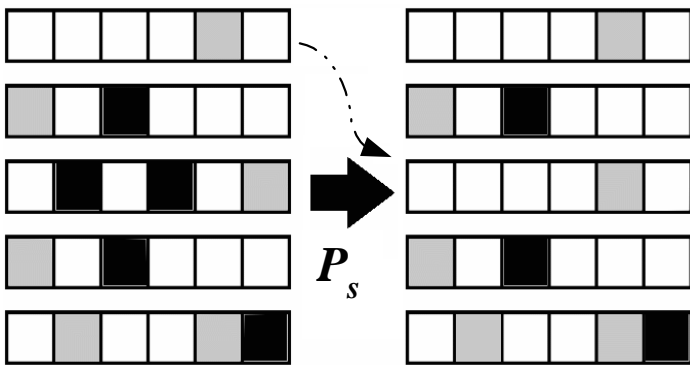
Objective function  
(Fitness function)

Encoding

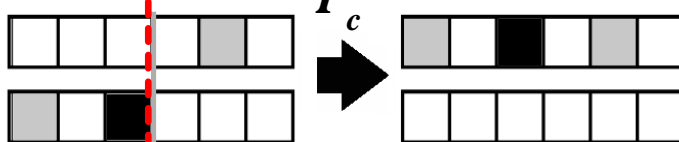
chromosome  
(individual)

gene

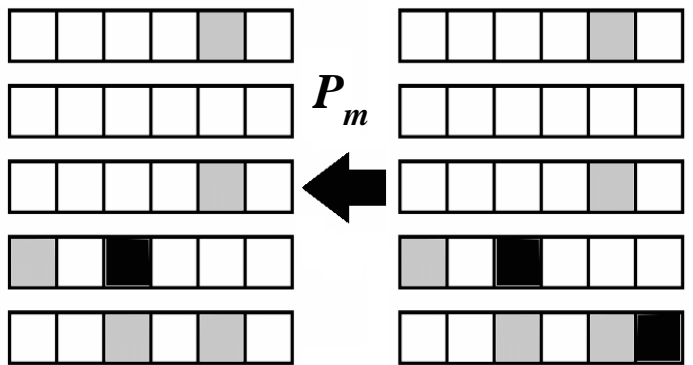
Selection



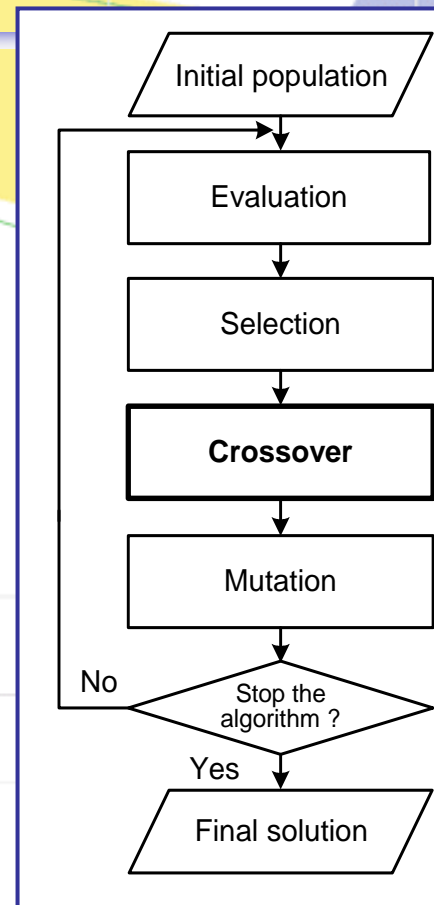
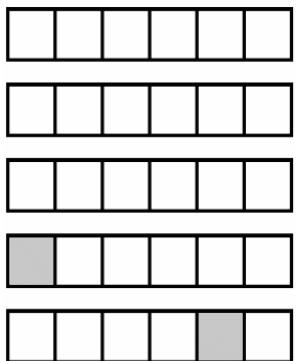
Crossover



Mutation



Results



# Intelligent Genetic Algorithm

## ➤ Drawback of conventional GA

- The performance of simple GA (SGA) suffer from the large number of optimization parameters of complicate problems
  - The search space exponentially grows when the number of optimization parameter increasing.
  - Fitness function only can evaluate an entire chromosome. It cannot evaluate the contribution of individual parameter.
  - It is difficult to generate high quality offspring using traditional crossover operators.



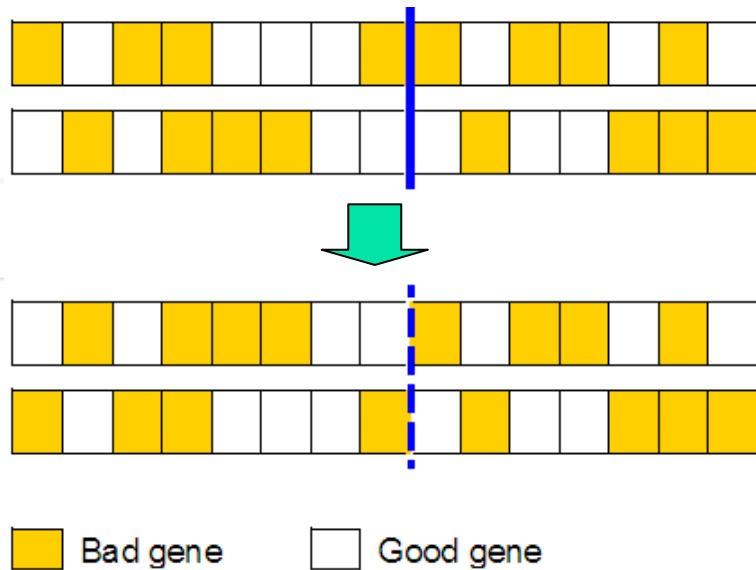
# ➤ Intelligent Genetic Algorithm

---

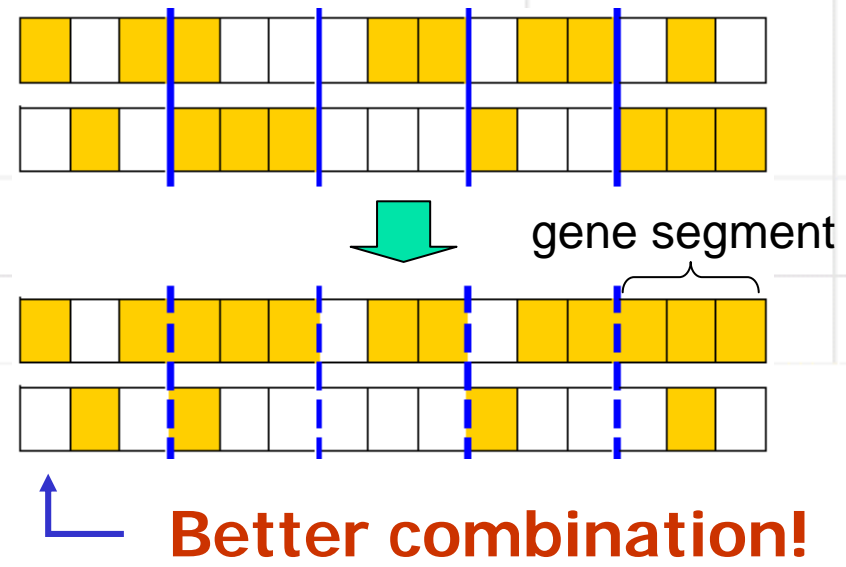
- Intelligent Genetic Algorithm (IGA)
  - Improve the crossover operator. It uses a systematic reasoning approach to evaluate the contribution of individual parameters to the fitness function, and recombine the better offspring according to the contributions.

# Intelligent Genetic Algorithm

## Conventional Random Crossover



## Intelligent Crossover





# ➤ Intelligent Genetic Algorithm

- Orthogonal Arrays (OA)
  - The orthogonal experimental design (OED) based on OA is a very efficient systematic reasoning mechanism which can obtain the optimal combination in linear computation time.
  - It has been widely applied to quality engineering, such as Taguchi Method
- Intelligent Crossover
  - Recombine high quality offspring from the parent chromosomes according to the contribution of each parameter to the fitness function using OED and factor analysis.



# Intelligent Genetic Algorithm

## $L_8(2^7)$ Orthogonal Array

Exp. no.	Factors							Evaluation value
	1	2	3	4	5	6	7	
1	1	1	1	1	1	1	1	$y_1$
2	1	1	1	2	2	2	2	$y_2$
3	1	2	2	1	1	2	2	$y_3$
4	1	2	2	2	2	1	1	$y_4$
5	2	1	2	1	2	1	2	$y_5$
6	2	1	2	2	1	2	1	$y_6$
7	2	2	1	1	2	2	1	$y_7$
8	2	2	1	2	1	1	2	$y_8$

# Intelligent Genetic Algorithm

## ➤ Intelligent Crossover

- Step 1: Use the first columns of OA  $L_{\beta}(2^{\beta}-1)$  where  $\beta=2^{\lceil \log_2(\beta+1) \rceil}$
- Step 2: Let levels 1 and 2 of factor  $j$  represent the  $j^{\text{th}}$  parameter of a chromosome coming from parents  $P_1$  and  $P_2$ , respectively.
- Step 3: Evaluate the fitness values  $y_t$  for experiment  $t$  where  $t=2, \dots, \beta$ . The value of  $y_1$  is the fitness of  $P_1$ .
- Step 4: Compute the main effect  $S_{jk}$

$$S_{jk} = \sum_{t=1}^{\alpha+1} y_t \cdot F_t, \quad j = 1, \dots, \alpha, \quad k = 1, 2$$

where  $F_t=1$  if the level of factor  $j$  of experiment  $t$  is  $k$ ;  
otherwise,  $F_t=0$ .

# Intelligent Genetic Algorithm

## ➤ Intelligent Crossover

- Step 5: Determine the better one of two levels of each factor. Select level 1 for the  $j^{\text{th}}$  factor if  $S_{j1} > S_{j2}$ . Otherwise, select level 2.
- Step 6: The chromosome of  $C_1$  is formed using the combination of the better genes from the derived corresponding parents.
- Step 7: The chromosome of  $C_2$  is formed similarly as  $C_1$ , except that the factor with the smallest main effect difference adopts the other level.
- Step 8: The best two individuals among  $P_1, P_2, C_1, C_2$ , and  $\beta-1$  combinations of OA are used as the final children  $C_1$  and  $C_2$  for elitist strategy.

# Intelligent Genetic Algorithm

## ➤ Further Reading

- Shinn-Ying Ho, L.-S. Shu, and J.-H. Chen, "Intelligent Evolutionary Algorithms for Large Parameter Optimization Problems," to be appeared in *IEEE Trans. Evolutionary Computation*, 2004.
- Shinn-Ying Ho, H.-M. Chen, S.-J. Ho, and T.-K. Chen, "Design of Accurate Classifiers with a Compact Fuzzy-Rule Base Using an Evolutionary Scatter Partition of Feature Space," *IEEE Trans. Systems, Man, and Cybernetics Part B*, 34/2, pp. 1031-1044, Apr 2004.

# Intelligent Genetic Algorithm

## ➤ Some Published Applications

- Data Mining
- VLSI Floor Planning
- Image Segmentation
- Polygonal Approximation
- Multi-Criteria Optimization
- Scheduling
- Multicast Routing
- System Identification
- Automatic Control System Design
- Fuzzy Neural Network Optimization

# The Evolutionary Design Approach

## ➤ Fuzzy Classifier

- Fuzzy theory has been successfully applied to classifier design.
  - Fuzzy neural network
  - Fuzzy  $k$ -NN classifier
  - Fuzzy rule-based classifier
- The linguistic rule representation of fuzzy rule-based classifier are more comprehensible than other fuzzy classifiers.

# The Evolutionary Design Approach

## ➤ Fuzzy Rule-Based Classifier

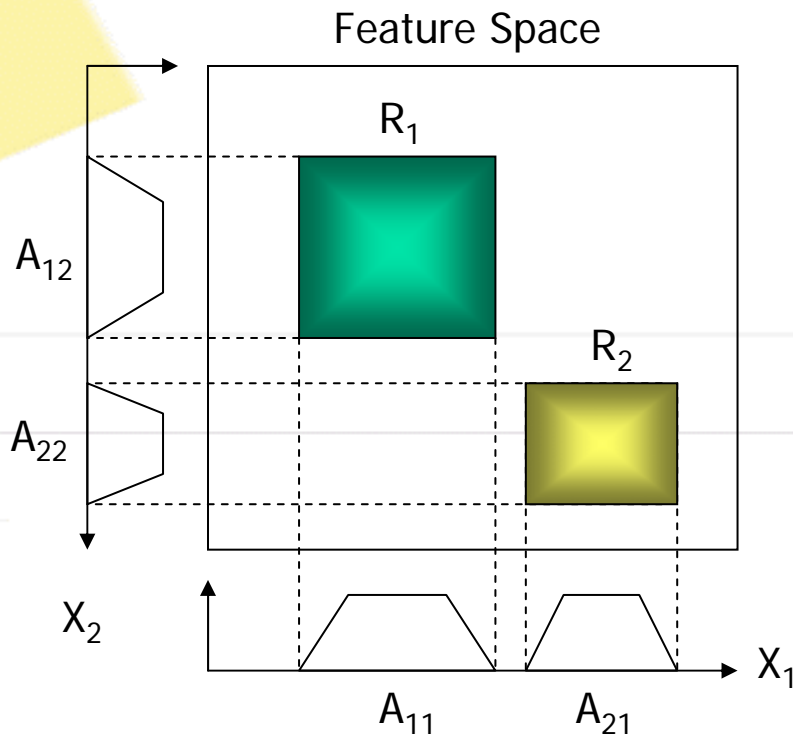
- Fuzzy Partition
  - The feature space is partitioned into regions using the membership functions.
  - Each region is assigned an appropriate rule.
- Fuzzy Reasoning
  - Calculate the fire strength of each rule, and
  - Determine the output by the voting result of rules



# The Evolutionary Design Approach

## Fuzzy Rule-Based Classifier

### 2D Example



Rule base:

$R_1$ : If  $x_1$  is  $A_{11}$  and  $x_2$  is  $A_{12}$  then  
Class 1 with CF=0.8

$R_2$ : If  $x_1$  is  $A_{21}$  and  $x_2$  is  $A_{22}$  then  
Class 2 with CF=0.5

$A_{ji}$  : Fuzzy Sets

$\mu_{A_{ji}}(x_i)$  : Membership functions

CF: Grade of certainty

# The Evolutionary Design Approach

## ➤ Fuzzy Rule-Based Classifier

- To full utilize the classification ability of fuzzy rule-based classifier, the optimization of the membership functions and the rule base is required.
- The fuzzy classifier design is formulated as a parameter optimization problem which is solved using evolutionary approaches.

# The Evolutionary Design Approach

## ➤ Fuzzy Rule-Based Classifier

- For real-world applications, the number of design parameters of fuzzy classifier would be large.
- The performance of conventional genetic algorithms suffer from the large parameter optimization problem.
- The proposed IGA can efficiently solve the difficult optimization problem.

# The Evolutionary Design Approach

## ➤ The Characteristics

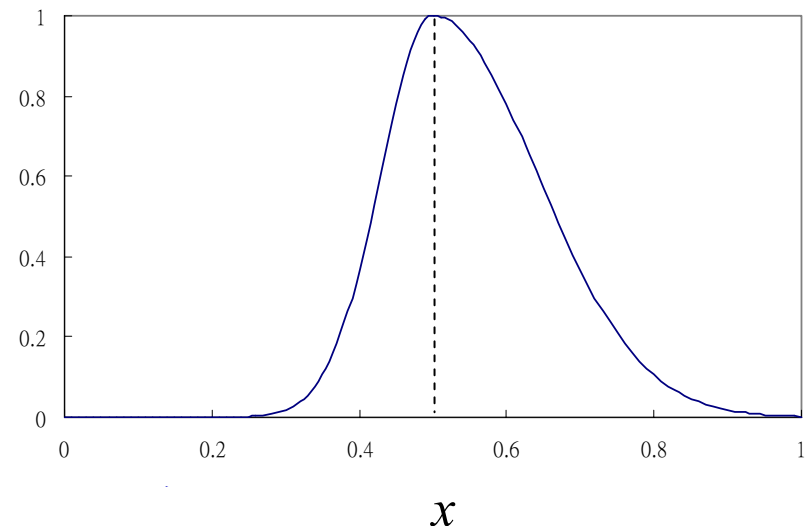
- The entire fuzzy rule base is encoded as a chromosome.
- Independent feature selection for each rule.
- Use clustering algorithms to obtain the initial location of membership function.
- Incorporate an efficient inheritance strategy with multi-stage evolution.

# The Evolutionary Design Approach

## ➤ Membership Functions

- A function represent the membership degree of an object  $x$  belonging to a fuzzy set A.
- We use an asymmetric Gaussian function as the membership function. The shape is determined by three parameters  $c$ ,  $\sigma_1$ ,  $\sigma_2$ .

$$\mu(x) = \begin{cases} e^{-\frac{(x-c)^2}{\sigma_1^2}} & \text{if } x < c \\ e^{-\frac{(x-c)^2}{\sigma_2^2}} & \text{if } x \geq c \end{cases}$$



# The Evolutionary Design Approach

## ➤ Membership Functions

- We also use a “*don't care*” membership function which can simplify the fuzzy rules.
  - The membership  $\mu(x)$  is always 1

- Example:

$R_1$ : If  $x_1$  is  $A_{11}$  then credit is good with  $CF=0.5$

$R_2$ : If  $x_1$  is  $A_{21}$  and  $x_2$  is  $A_{22}$  then credit is bad with  $CF=0.8$

where  $R_1$  has an implicit condition:  $x_2$  is “*don't care*”

# The Evolutionary Design Approach

## ➤ Fuzzy Reasoning

### ■ Calculation of fire strength

- Let  $x_p = (x_{p1}, x_{p2}, \dots, x_{pn})$  be the vector of input variables, the fire strength  $\mu_{R_j}(x_p)$  (the membership degree of  $x_p$  belong to the rule) of  $R_j$  is determined as follows:

$$\mu_{R_j}(x_p) = \mu_{A_{j1}}(x_{p1}) \cdot \dots \cdot \mu_{A_{jn}}(x_{pn}) = \prod_{i=1}^n \mu_{A_{ji}}(x_{pi})$$

# The Evolutionary Design Approach

## ➤ Fuzzy Reasoning

- Determine the output class of  $x_p$ 
  - Step 1: Calculate the scores  $S_{\text{class } v}$ ,  $v=0,1$  (good credit and bad credit)

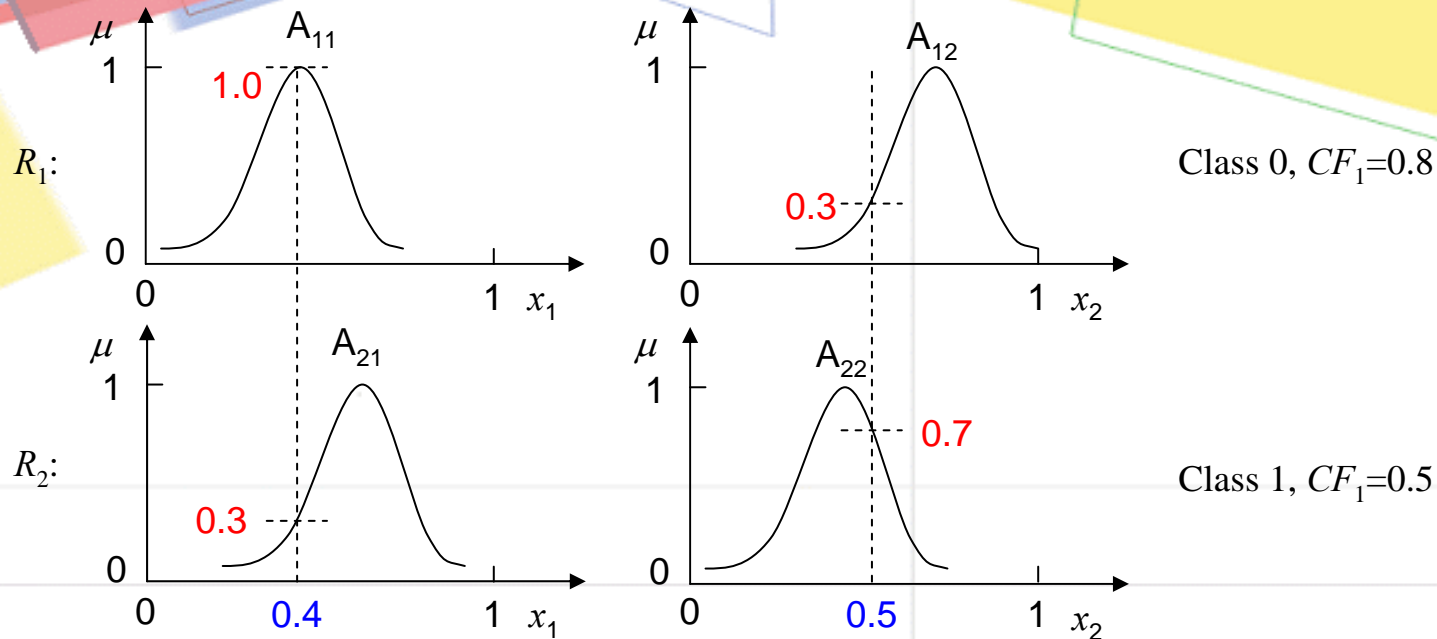
$$S_{\text{class } v} = \sum_{\substack{C_j=v \\ j=1,\dots,N}} \mu_{R_j}(x_p) \cdot CF_j$$

- Step 2: Classify  $x_p$  to the class with the highest score  $S_{\text{class } v}$



# The Evolutionary Design Approach

## Fuzzy Reasoning



Given  $x_p=(0.4, 0.5)$ , then

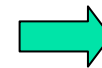
$$\mu_{R_1}(x_p) = 1.0 \times 0.3 = 0.30$$

$$\mu_{R_2}(x_p) = 0.3 \times 0.7 = 0.21$$



$$S_{\text{class0}} = 0.30 \times 0.8 = 0.240$$

$$S_{\text{class1}} = 0.21 \times 0.5 = 0.105$$



$x_p$  is Class 0  
(Score: 0.135)

# The Evolutionary Design Approach

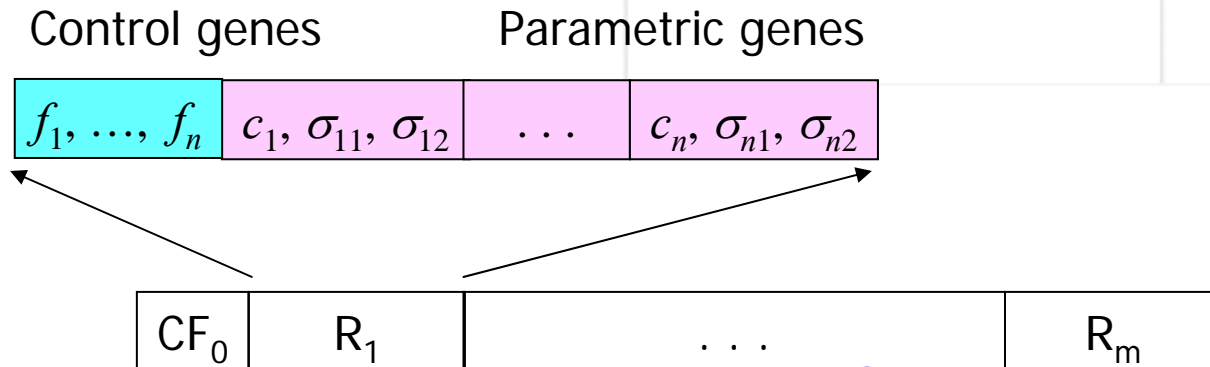
## ➤ Chromosome Representation

- Rule  $R_0$ 
  - All the antecedent conditions are “*don't care*”
  - Output class is 0
  - $CF_0$  is evolved by IGA
  
- Rules  $R_1, \dots, R_m$ 
  - The antecedent conditions are evolved by IGA
  - Output classes are 1
  - $CF_j$  is determined by heuristic.

# The Evolutionary Design Approach

## ➤ Chromosome Representation

- The control genes perform feature selection for rules.
  - If  $f_i = 0$ , the corresponding membership function of  $x_i$  is “*don't care*”
  - If  $f_i = 1$ , the corresponding membership function of  $x_i$  is Gauss2( $c_i, \sigma_{i1}, \sigma_{i2}$ )



# The Evolutionary Design Approach

## ➤ Grade of Certainty

- Use the method proposed by H. Ishibuchi
  - Step 1: Calculate the fire strength of  $R_j$  for each training sample  $x_p$
  - Step 2: Calculate the sum of strengths  $\beta_0, \beta_1$

$$\beta_0 = \sum_{x_p \in \text{class0}} \mu_{R_j}(x_p) \quad \beta_1 = \sum_{x_p \in \text{class1}} \mu_{R_j}(x_p)$$

- Step 3:  $CF$  is defined as

$$CF = \max\left(\frac{\beta_0 - \beta_1}{\beta_0 + \beta_1}, 0\right)$$

# The Evolutionary Design Approach

## ➤ Fitness Function

- Minimizing the misclassification cost is equal to maximizing the profit of correct classification.
- The fitness functions is defined as

$$\max F(FC) = C_0 \cdot NCP_0 + C_1 \cdot NCP_1,$$

*FC*: the decoded fuzzy classifier

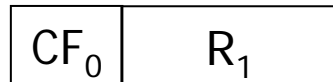
*NCP*<sub>0</sub>, *NCP*<sub>1</sub>: the number of correct classified patterns of class 0 and class 1

*C*<sub>0</sub>, *C*<sub>1</sub>: the weight values of cost,

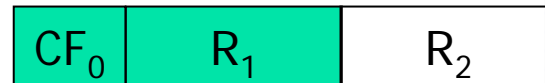
# The Evolutionary Design Approach

## ➤ Inheritance Strategy

Stage 1:

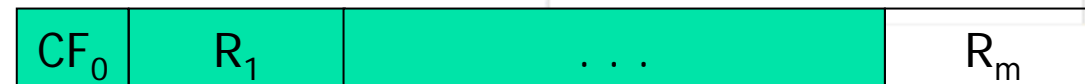


Stage 2:



⋮

Stage m:



**Inherit**

# The Evolutionary Design Approach

## ➤ The Procedural

- Step 1: Let the number of rules of class 1  $N_{r_1}=1$
- Step 2: Partitioning the class 1 samples into  $N_{r_1}$  clusters using  $k$ -means.
- Step 3: Initialization of population
  - First half: Center parameters  $c_j$  of membership functions are the means of clusters. Other parameters are assigned a random value.
  - Others:  $CF_0$  and the parameters of the first  $N_{r_1}-1$  rules direct inherit from the previous evolution stage, and the parameters of the  $N_{r_1}$ -th rule are assigned a random value. If no previous stage, the initialization is the same as the first half.

# The Evolutionary Design Approach

## ➤ The Procedural

- Step 4: Optimize the parameters of current stage using IGA
- Step 5: Add the optimal solution to the Pareto set.
- Step 6: If  $N_{r1}$  reaches predefined  $N_{r1\_max}$ , stop the multi-stage evolution.
- Step 7: Otherwise, let  $N_{r1} = N_{r1} + 1$  and go to step 2.



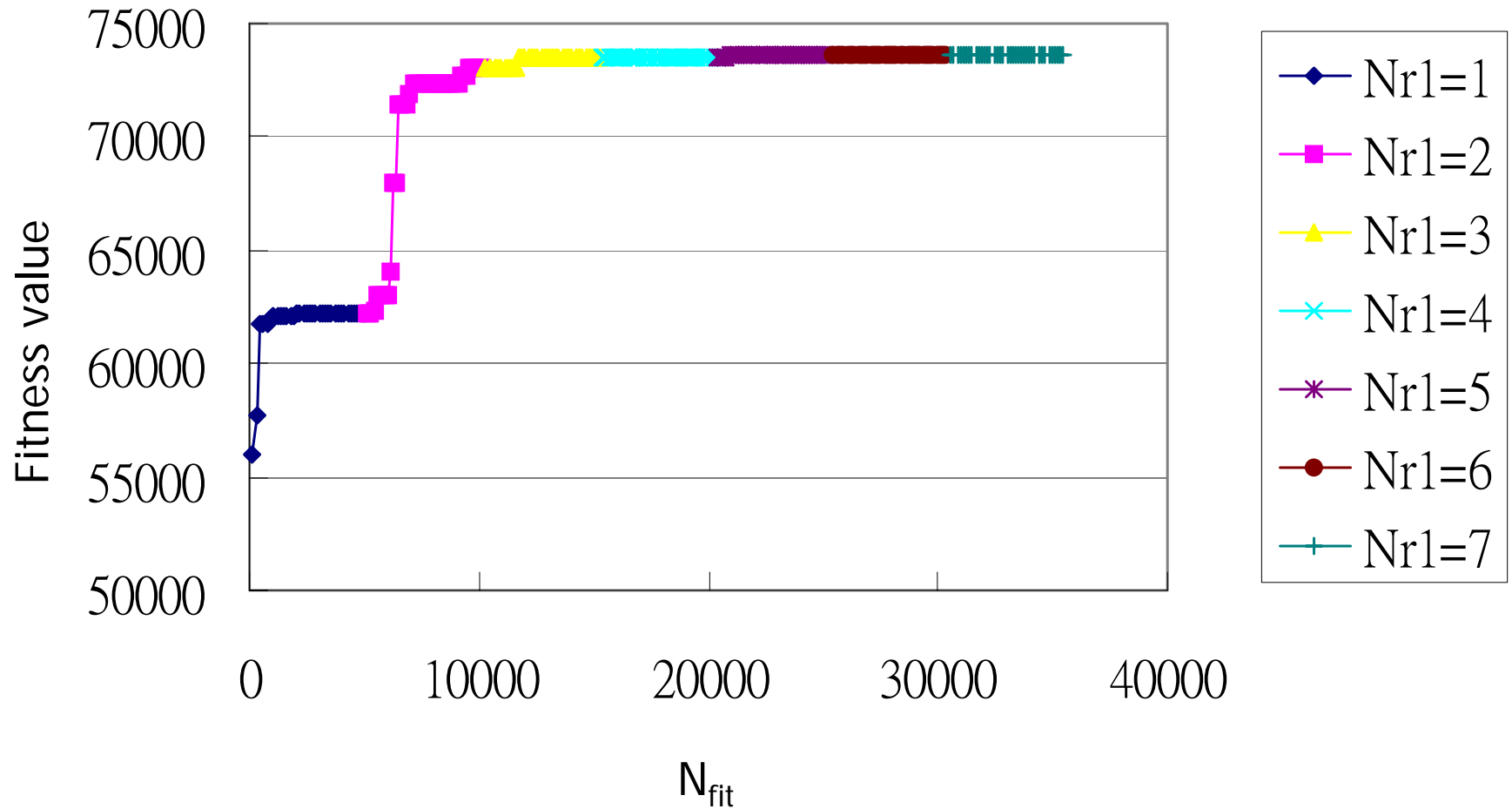
# Experimental Result

## ➤ Data Set and Parameter Setting

- Credit Scoring Data Set
  - Number of features: 28
  - Training samples: 60,000
  - Test samples: 6,638
- Weight values of cost:  $C_0=1$ ,  $C_1=5$
- The parameter setting of IGA
  - $N_{pop}=20$
  - $P_c=0.8$
  - $P_m=0.05$
  - Max. number of function evaluation: 5000 times per stage

# Experimental Result

## Convergence of Multi-Stage Evolution



# Experimental Result

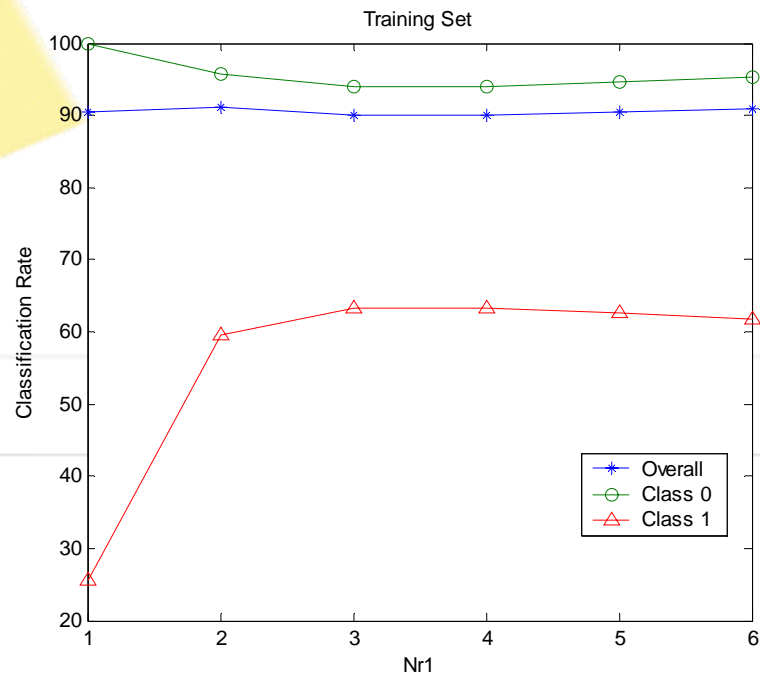
## ➤ Classification Accuracy

Nr1	Training			Test		
	overall	class 0	class 1	overall	class 0	class 1
1	90.42	100.00	25.65	88.51	100.00	29.22
2	91.09	95.76	59.48	90.98	96.38	63.08
3	90.01	93.97	63.21	89.80	94.32	66.51
4	90.01	93.97	63.21	89.80	94.32	66.51
5	90.42	94.53	62.60	90.09	94.73	66.14
6	90.96	95.28	61.72	90.84	95.90	64.75

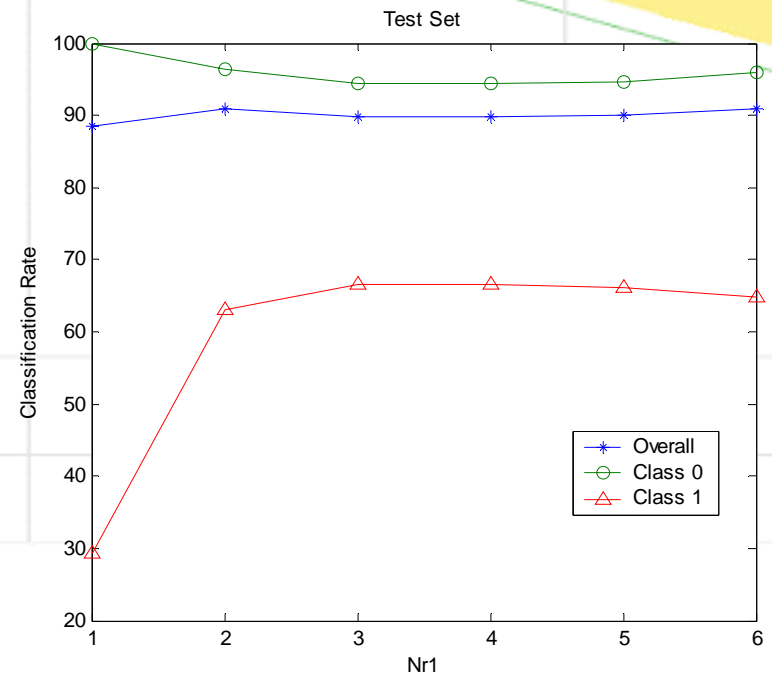
# Experimental Result

## Classification Accuracy

### Training Samples



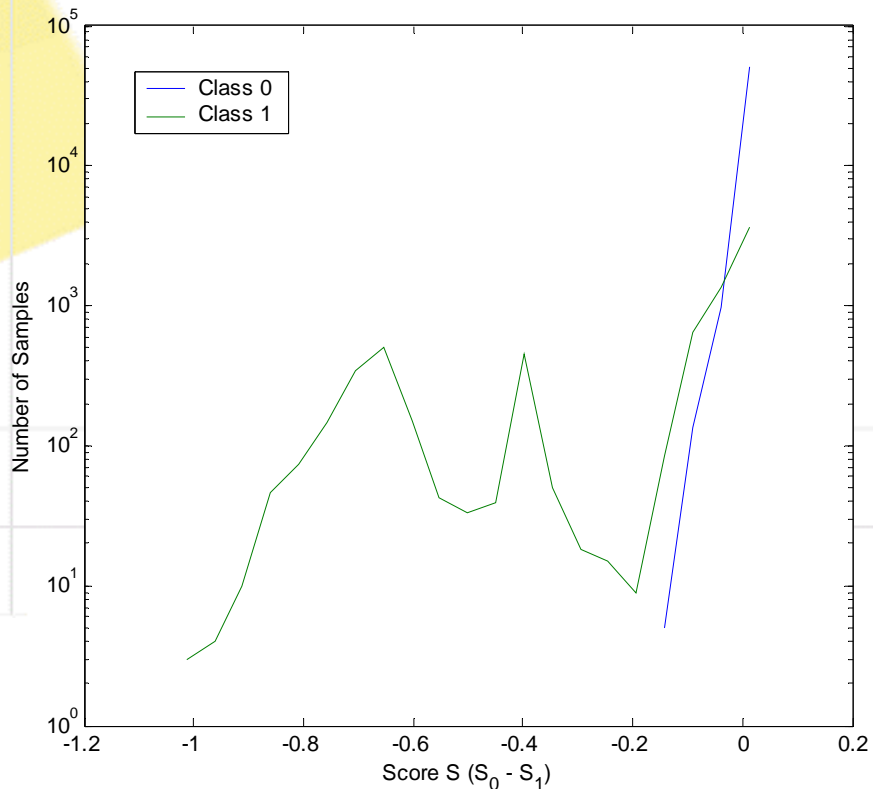
### Test Samples



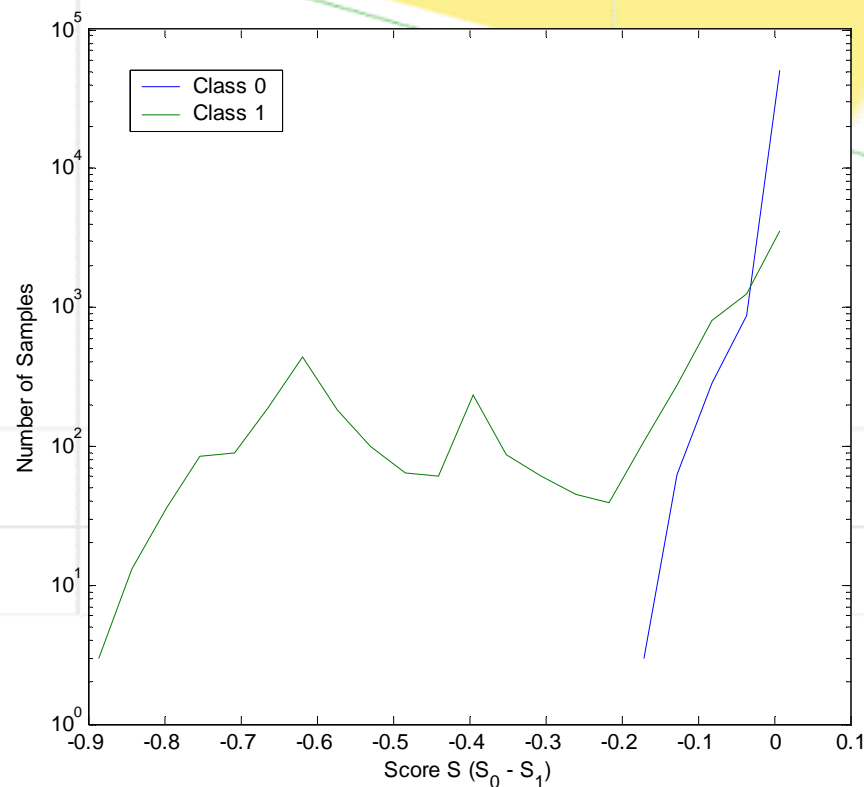
# Experimental Result

## Histogram of Scores (training samples)

$N_{r1}=2$



$N_{r1}=5$

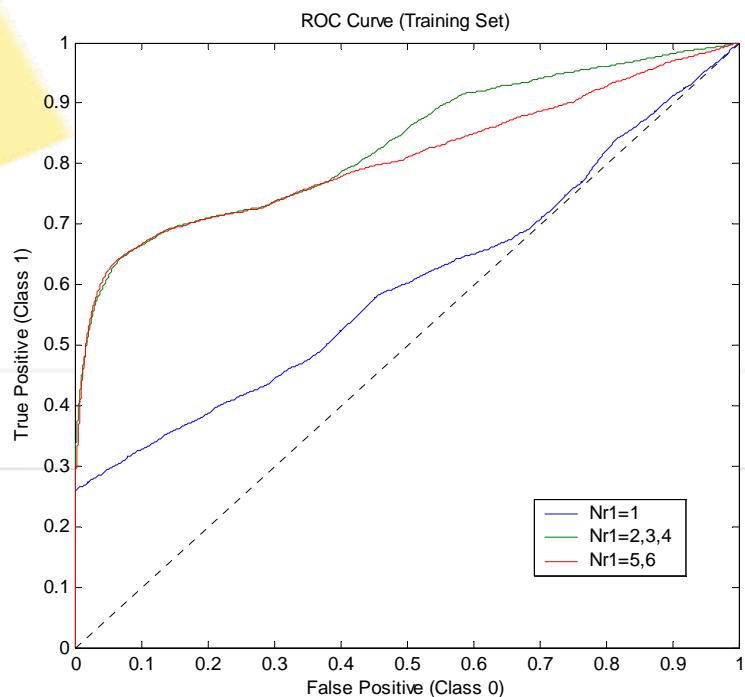


(semi-log scale for Y-axis)

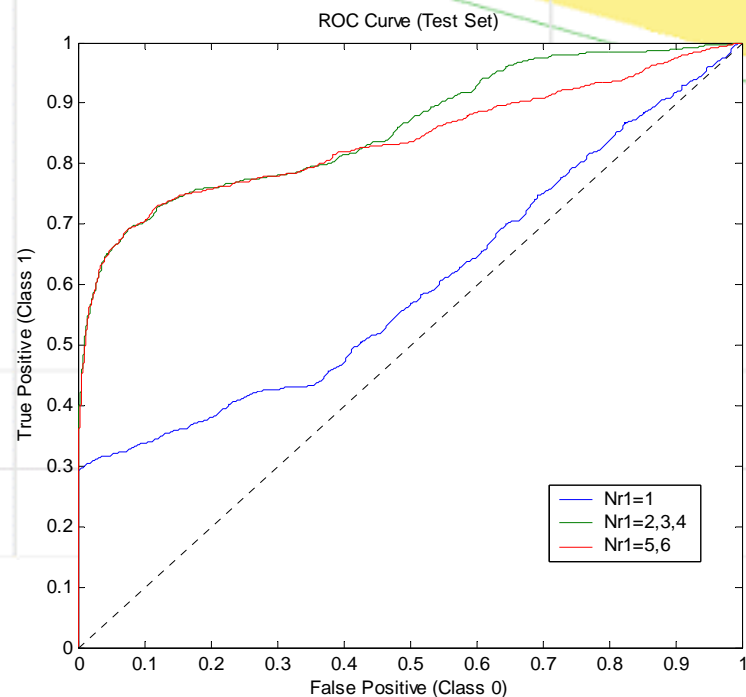
# Experimental Result

## ROC Curves

### Training Samples



### Test Samples





## Discussions

- The proposed approach can generate a set of solutions. The choice of solution is depended on user's preference.
- The proposed approach can generate quantitative scores for input patterns.
- The histogram of scores is similar to the probability density model.



## Conclusions

- Using machine intelligence to establish accurate credit scoring model would be popular in the near future.
- The proposed evolutionary approach can obtain accurate and interpretable credit scoring models.
- It can be further extended to design credit rating systems and probability models.





**~ Thank You ~**