

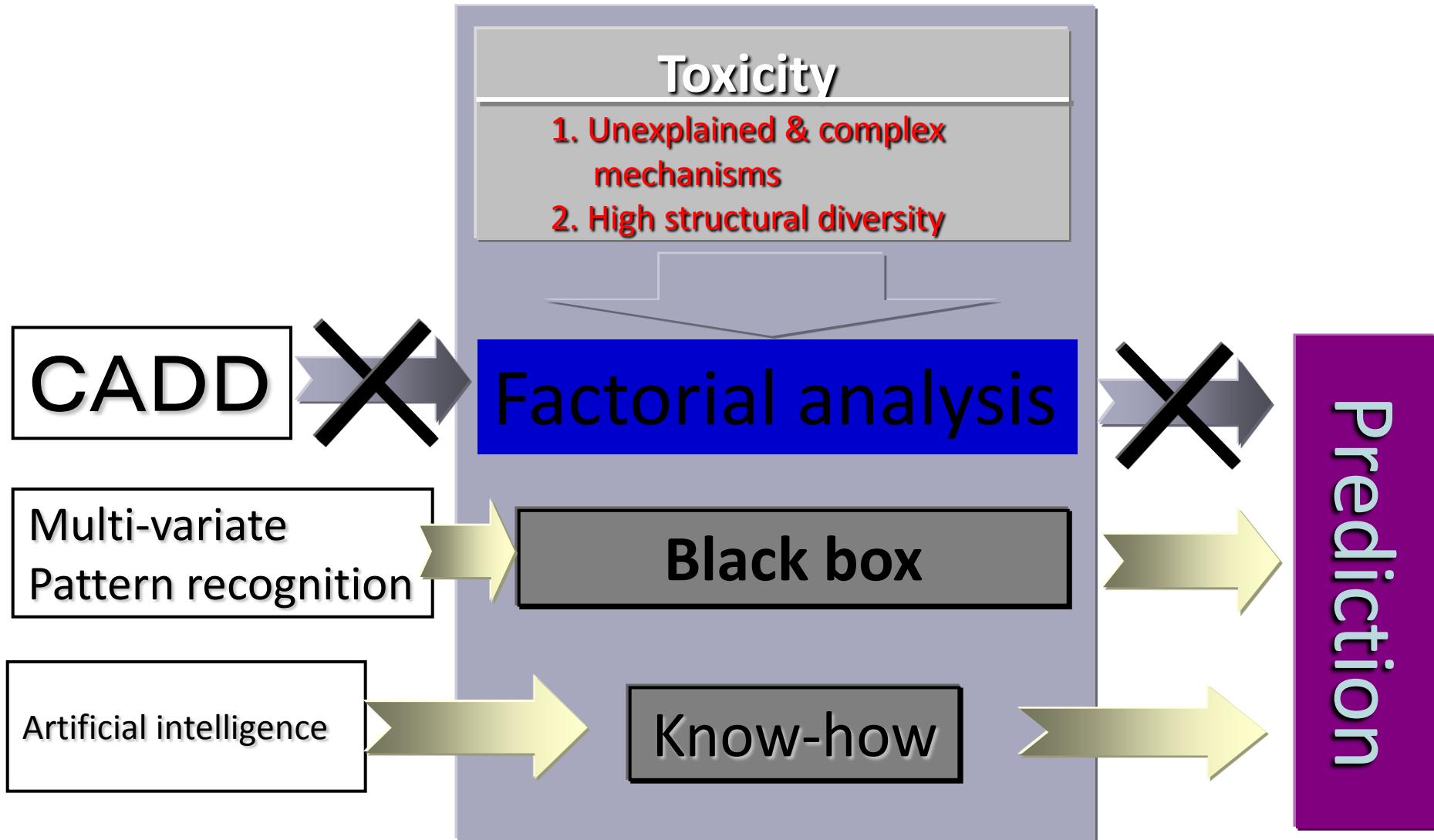
PI-111



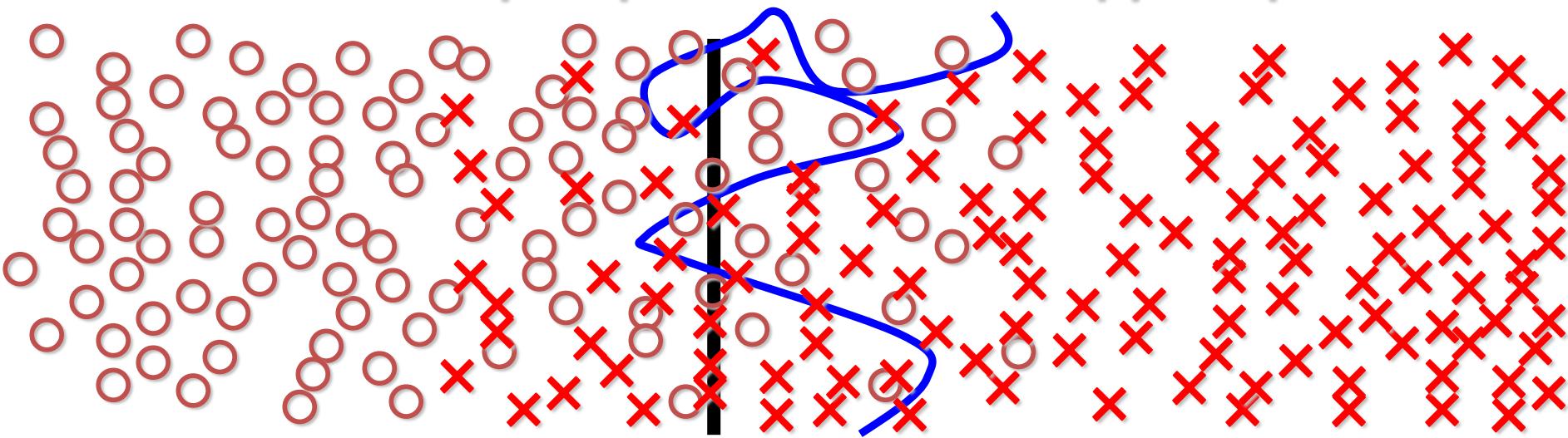
A NEW CLASSIFICATION METHOD SUITABLE FOR TOXICITY SCREENING OF CHEMICAL COMPOUNDS

Kohtaro Yuta
In Silico Data Ltd.
E-Mail : k-yuta@insilicodata.com
<http://www.insilicodata.com>

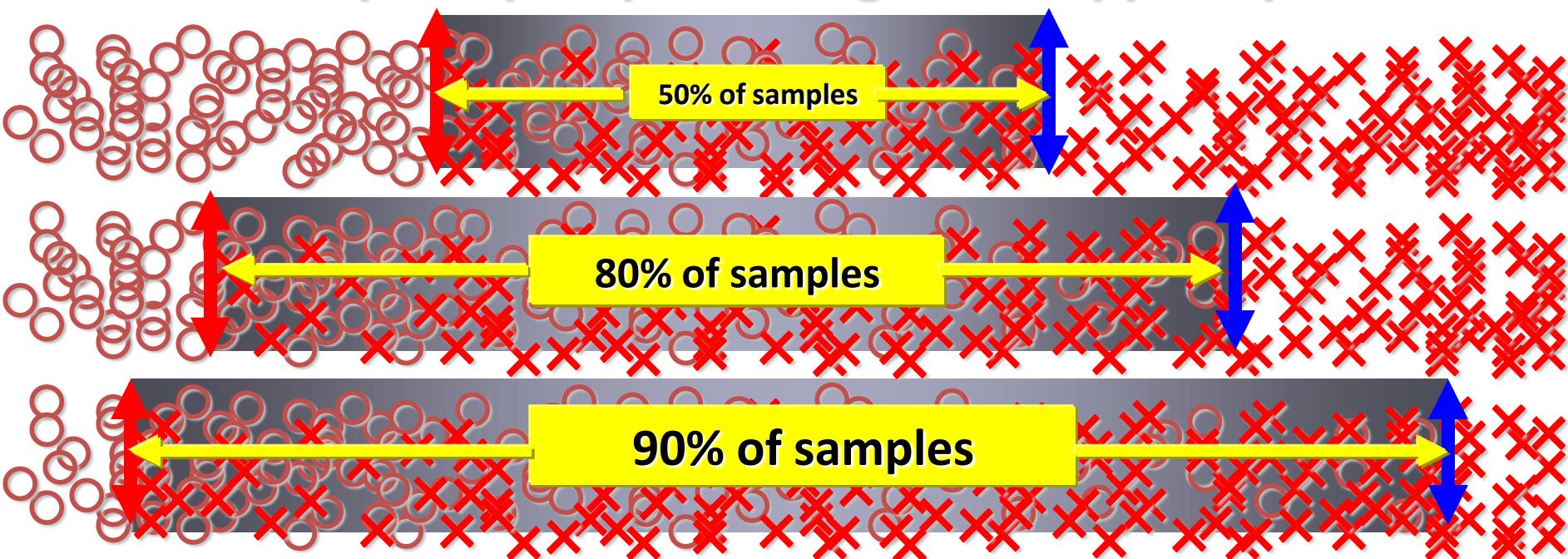
Why PR(Pattern Recognition) for toxicity screening



Normal sample space : small overlapped space



Toxicity sample space : large overlapped space



Perfect(100%) classification of Ames test 6965 pos/neg sample set

K-step Yard sampling method **KY-method**



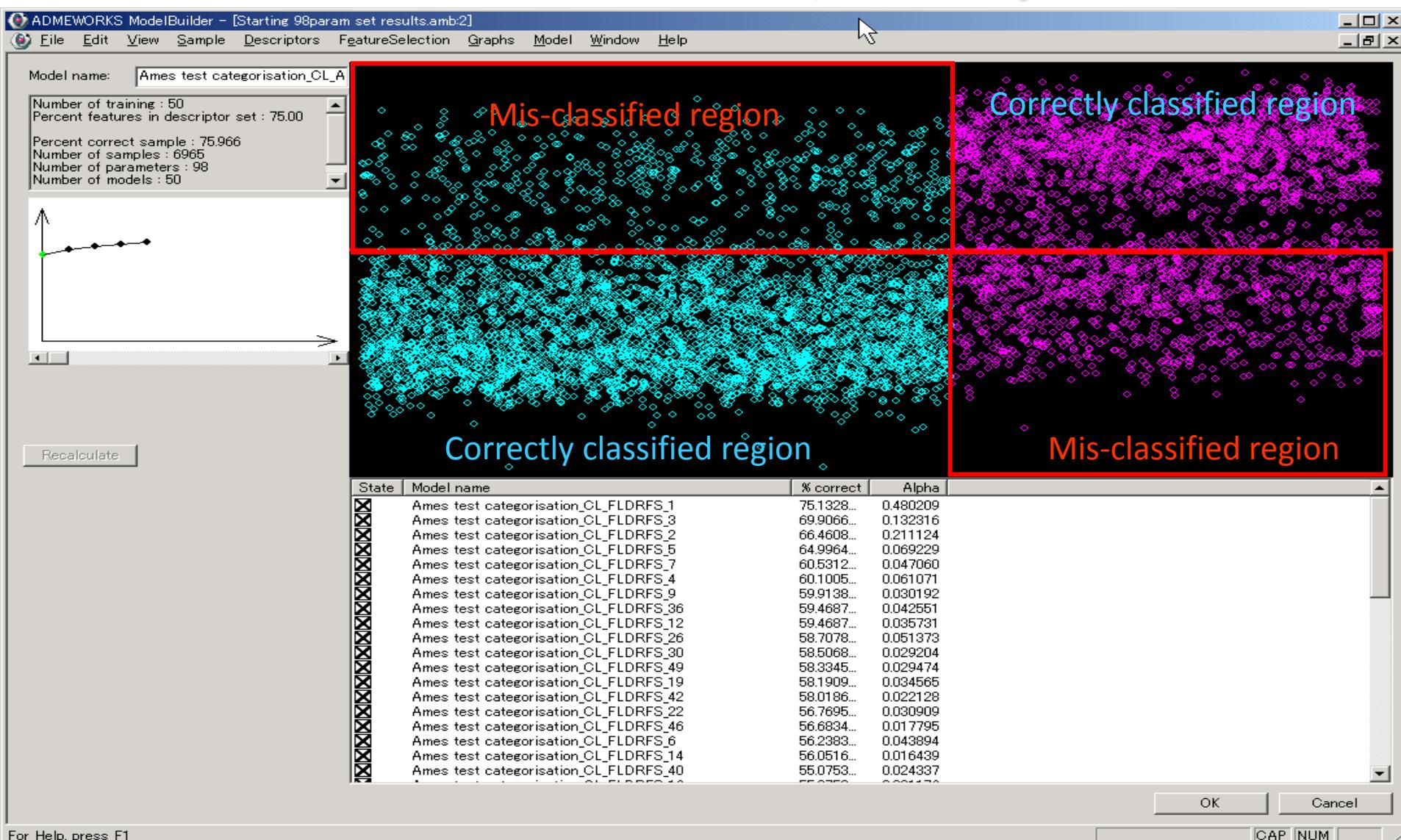
The most powerful and advanced data analysis method



The most difficult classification problem
**6,965 sample of Ames test were,
Classified perfectly**

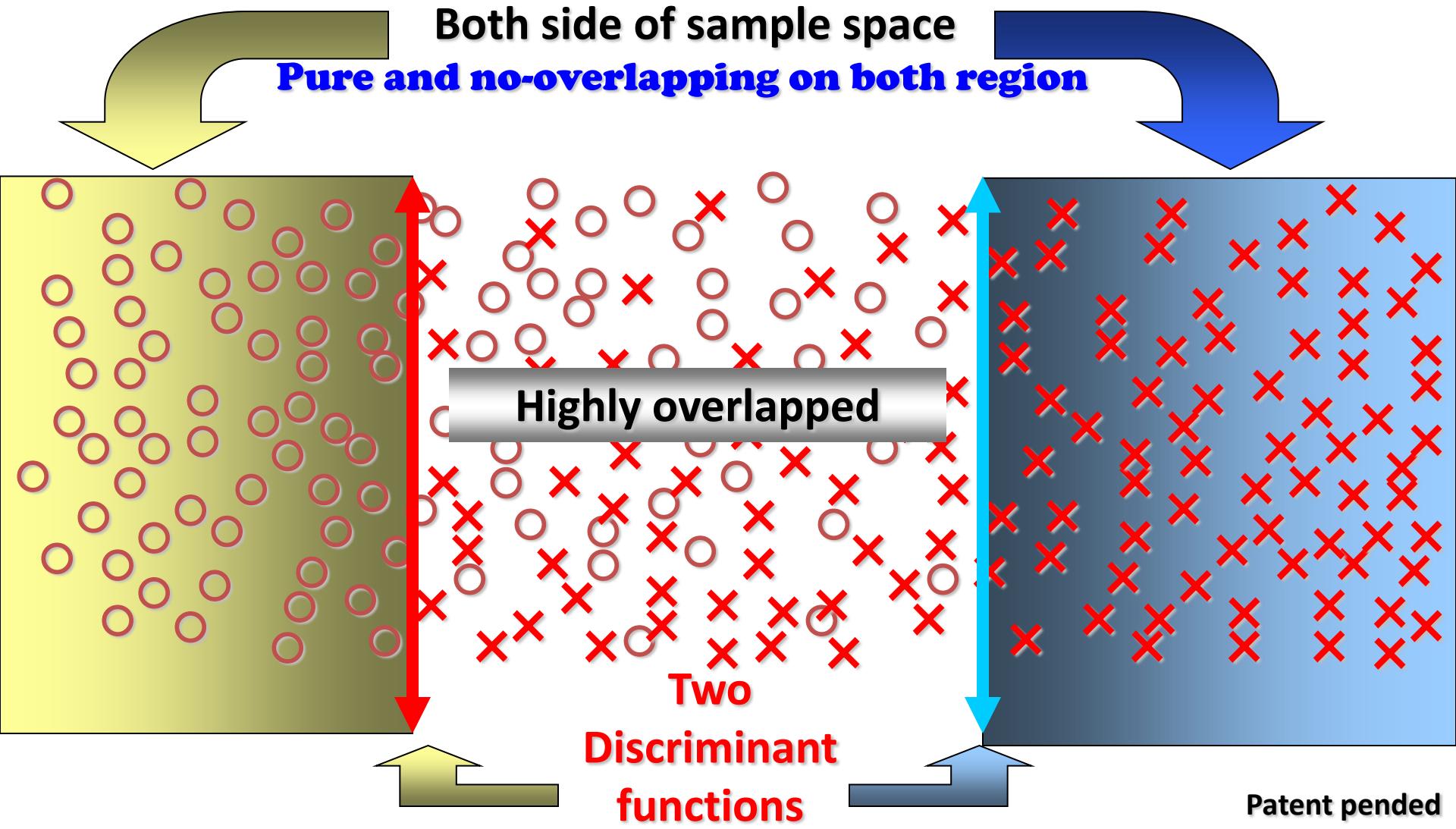
Classification Result by AdaBoost

77.24% of Ames test 6,965 samples



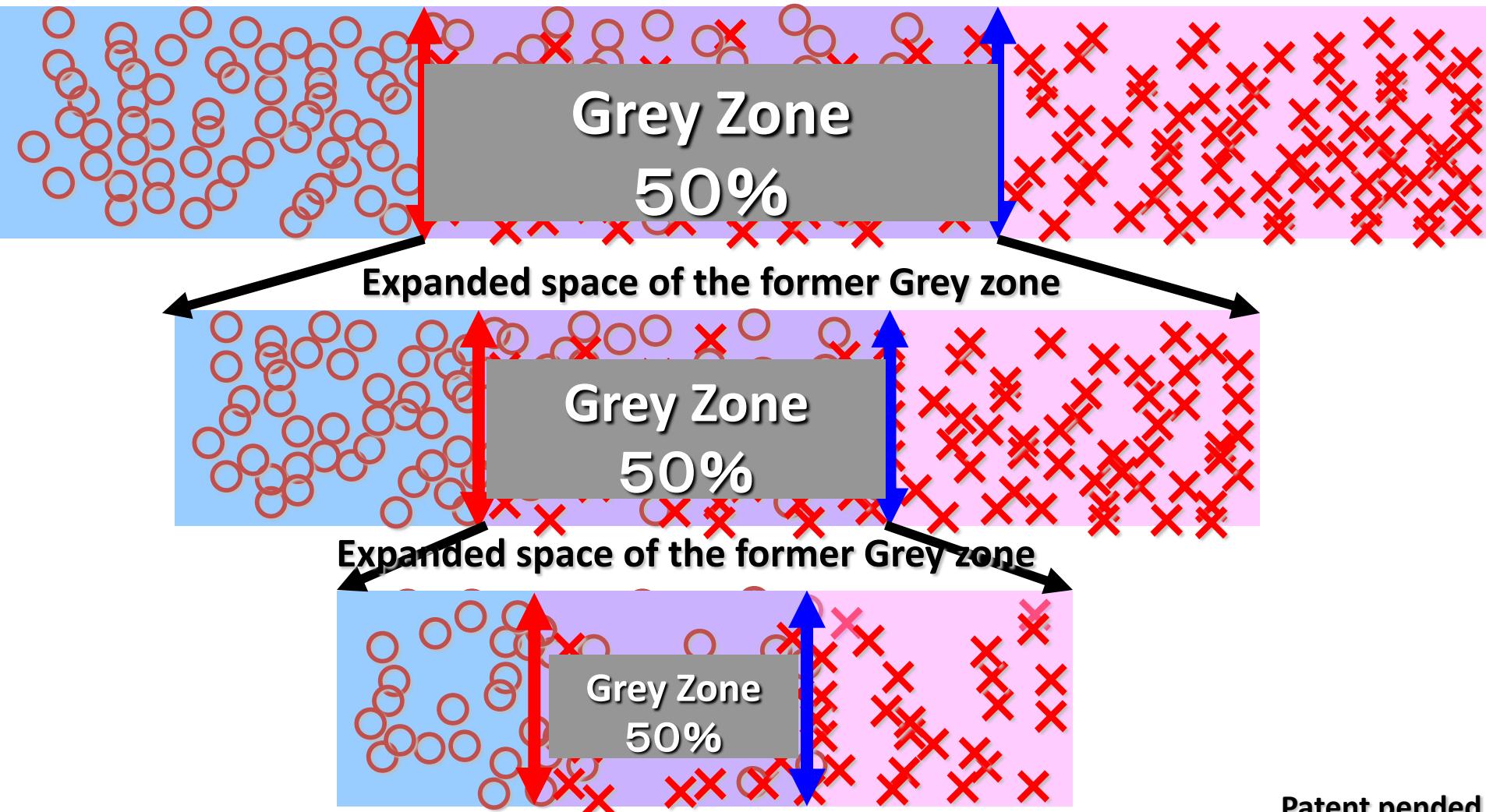
First basic concept of KY method

Spatial region on sample space

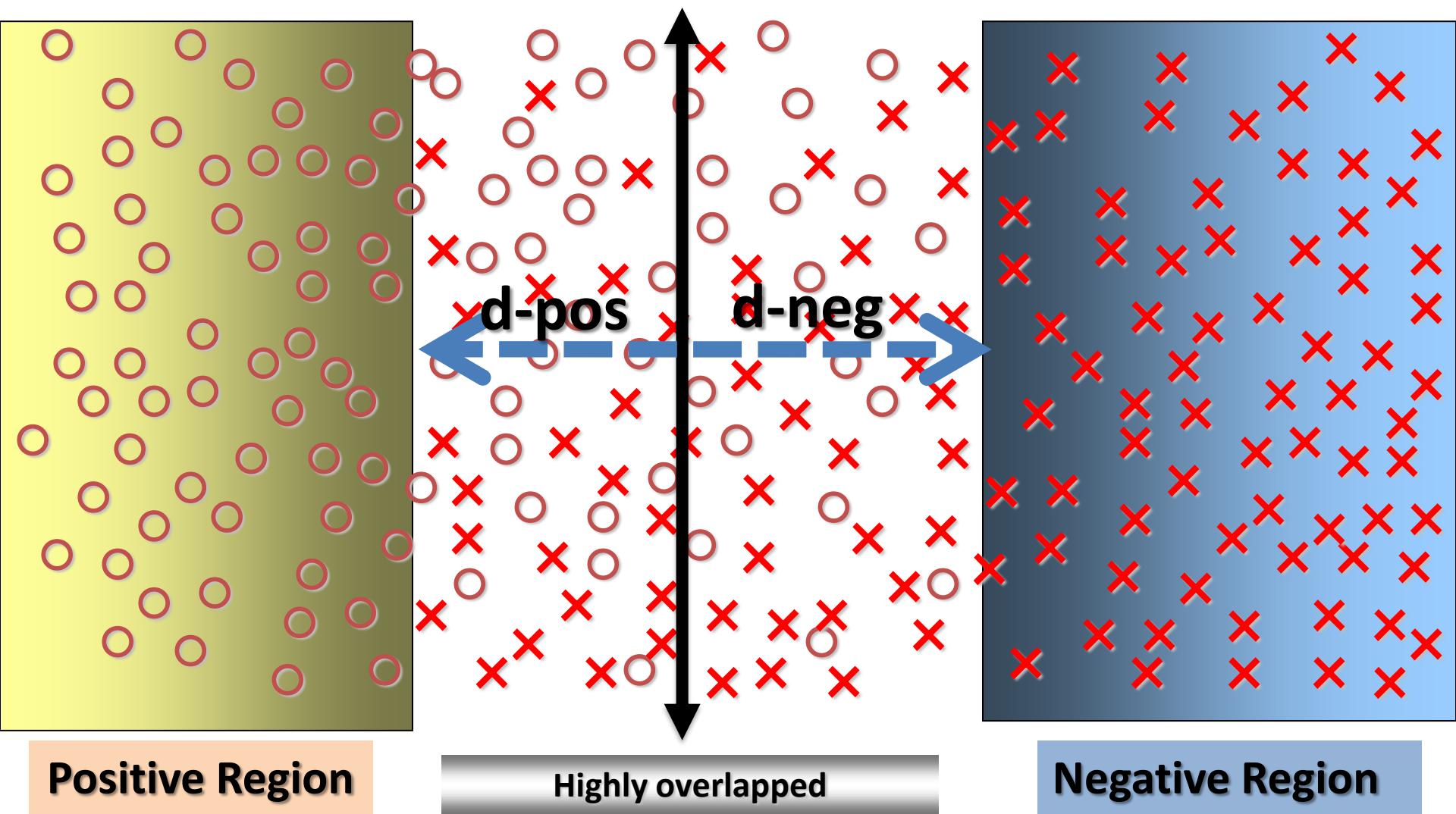


Second basic concept of KY method

Multi-steps for 100% classification



New approach to the “KY method” by one discriminant function



A series of KY methods

Discriminant Analysis	Fitting
Two model KY	KY Fitting with DA
Single model KY	KY Fitting with no DA
Model free KY	Model free KY Fitting

*Always carry perfect classification

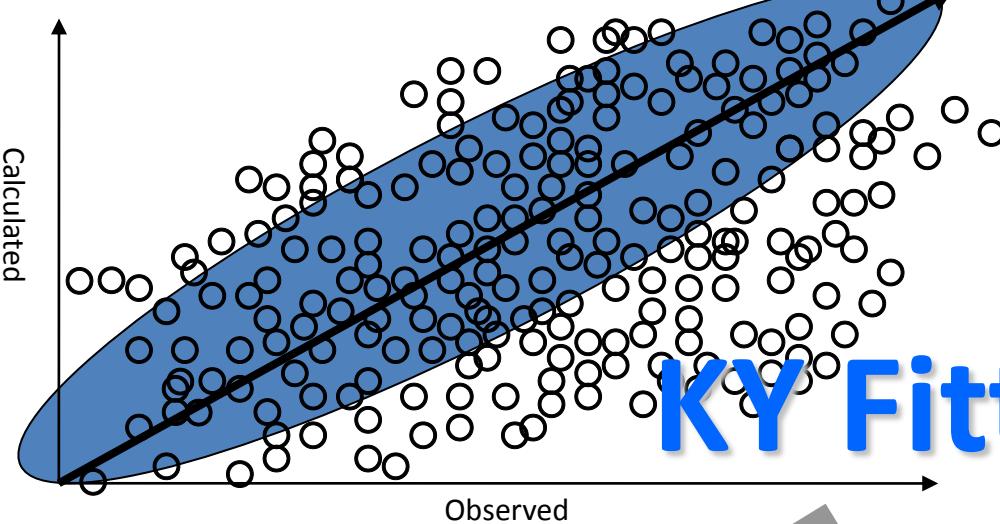
*Always high coefficient of determination

Discriminant Analysis	Fitting
Tailor-made Modeling for DA	Tailor-made Modeling for Fitting

*Always carry high prediction ratio

All methods were Patent Pended

Starting large sample set



Inner samples (G_i)

Calculated

$R=0.9$

Observed

KY Fitting

with DA

Move from Outer space to Inner space

Calculated

Observe

$$Y(\text{all}) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \text{Const.}$$

New sample space
generated from
outer space samples

Calculated

Observed

Patent pending

◆ KY method for fitting methods (Will be soon coming)

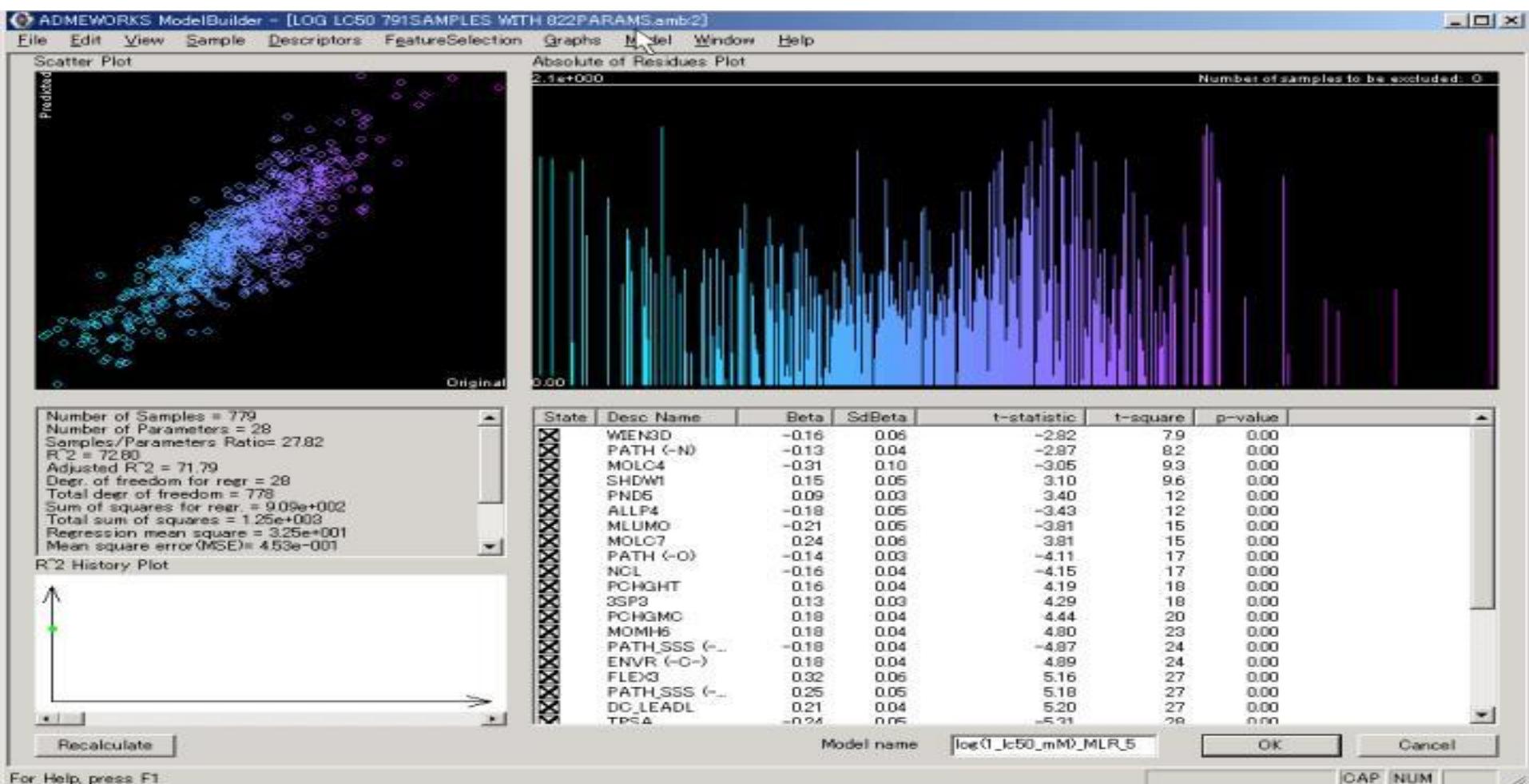
Fish: 96 hours LC50、Number of samples: 791、 $\text{Log}(1/\text{LC50_Mm})$ (Max/Min) : 6.376 / -2.963

◆ Data analysis by ordinal linear regression

Step1: **Inner** sample set

Number of samples: 779, Number of used parameters: 28, Confidance ratio: 27.8

R2: 72.8, R: 85.3, F-value: 71.7, CV: 69.6

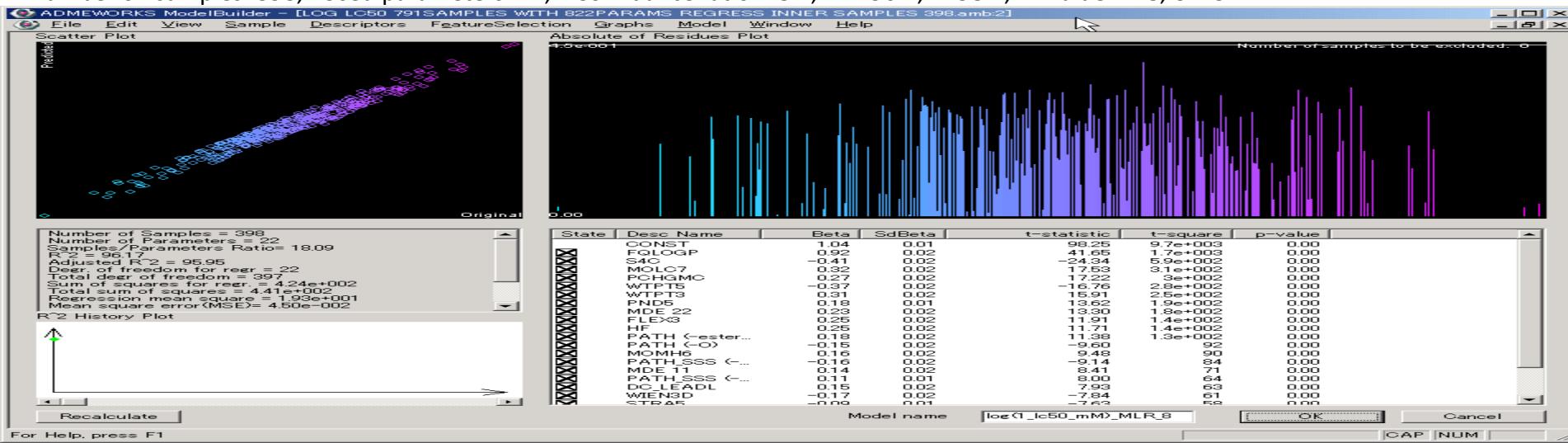


◇Fitting KY method Step1 (Inner sample set)

Patent pending

Step1: **Inner** sample set

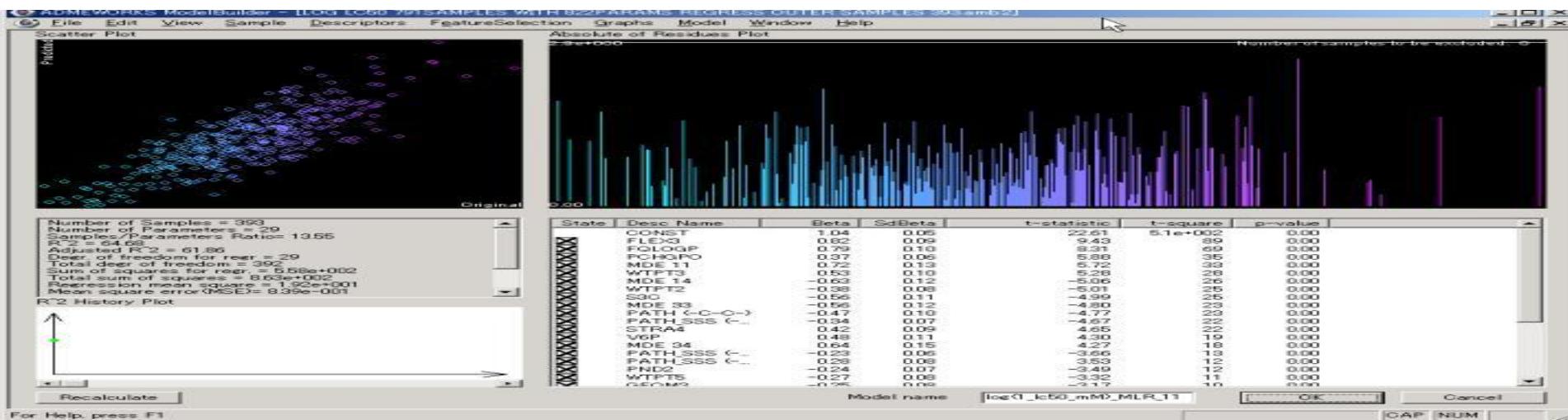
Number of samples: 398, Used parameters: 22, Confidance ratio: 18.1, R2: 96.2, R: 98.1, F-value: 428, CV: 94.4



◇Fitting KY method Step1 (Outer sample set)

Step1: **Outer** sample set

Number of samples: 393, Used parameters: 29, Confidance ratio: 13.6, R2 : 64.7, R: 80.4, F-value: 22.9, CV: 57.5



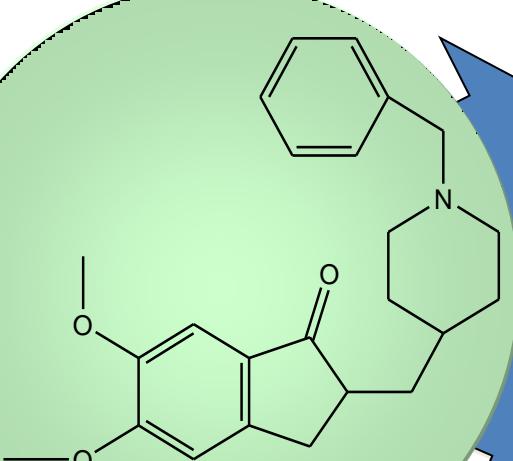
Drug properties and compound structure

ADME
properties

Pharmacological
activity

Physicochemical
properties

Toxicity



All properties are fixed when the structure is determined.

There are no relations between any two properties.

All properties must be optimized for developing drugs.

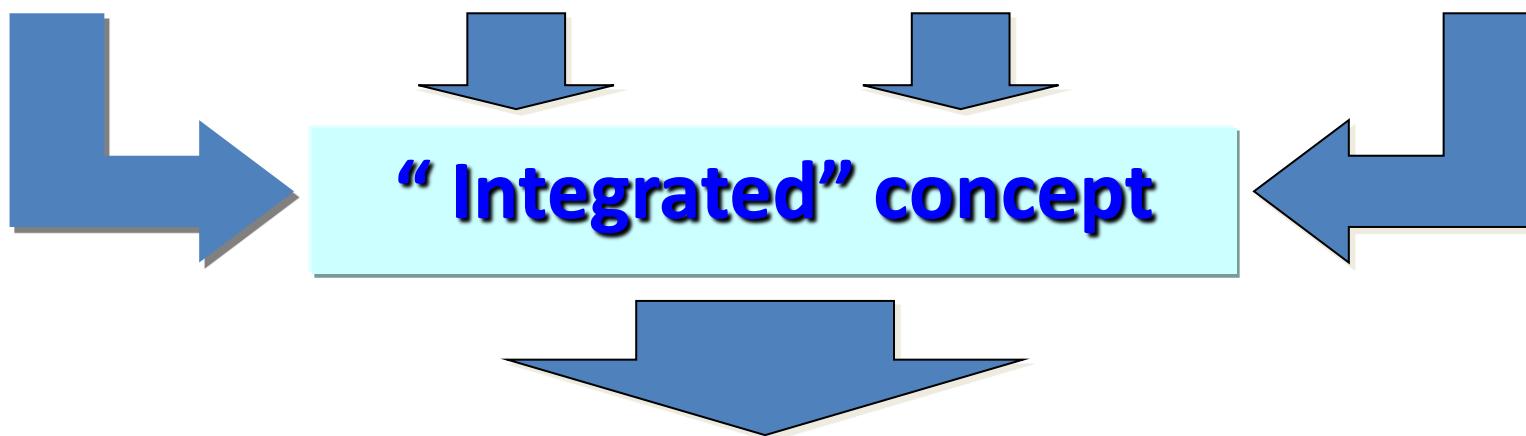
“Integrated” concept In Silico Data

Miracles by the KY-methods

for drug development

Activity + ADME + Toxicity + Property

All drug properties shall be considered at the same time



“Integrated” in silico screening & drug design

Flow of the “Parallel & One Step” D.D.

“Parallel & One Step” D.D.

In Silico prediction

Activity

ADME

Toxicity

Property

Wet Experiment

SYNTHESIS

Confirmation

TEST

Phase
I/II/III

