ENSEMBLE METHOD

BOOSTING, STACKING, BAGGING
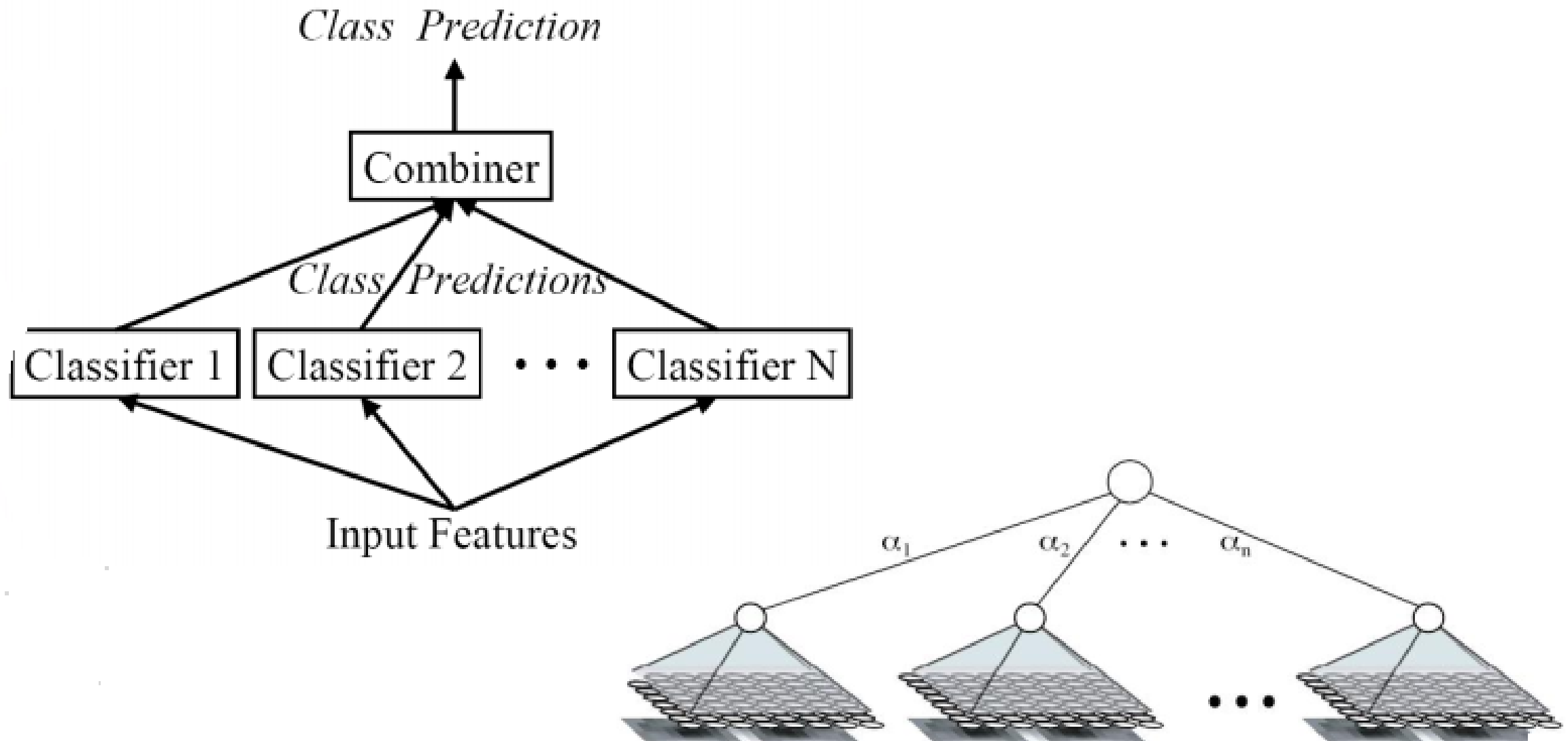

LE BICH PHUONG - HUMG

MACHINE LEARNING

CLASSICAL LEARNING

UNSUPERVISED

Clustering
- DBSCAN
- K-Means
- Agglomerative
- Mean-Shift
- Fuzzy C-Means

Pattern search
- Euclat
- Apriori
- FP-Growth

DIMENSION REDUCTION (generalization)
- t-SNE
- PCA
- LSA
- SVD
- LDA

SUPERVISED

Classification
- Naive Bayes
- K-NN
- SVM
- Decision Trees
- Logistic Regression

Regression
- Linear Regression
- Polynomial Regression
- Ridge/Lasso Regression

REINFORCEMENT LEARNING
- Genetic Algorithm
- Q-Learning
- SARSA
- Deep Q-Network (DQN)
- A3C

ENSEMBLE METHODS
- Stacking
- Bagging
- Random Fores
- Boosting
- XGBoost
- AdaBoost
- LightGBM
- CatBoost

NEURAL NETS AND DEEP LEARNING

Convolutional Neural Networks (CNN)
- DCNN

Recurrent Neural Networks (RNN)
- LSM
- LSTM
- GRU

Generative Adversarial Networks (GAN)

Autoencoders
- seq2seq

Perceptrons (MLP)

Real World Scenarios

Combination of Classifiers

$$d : \mathbb{R}^n \to \{0,1\}$$
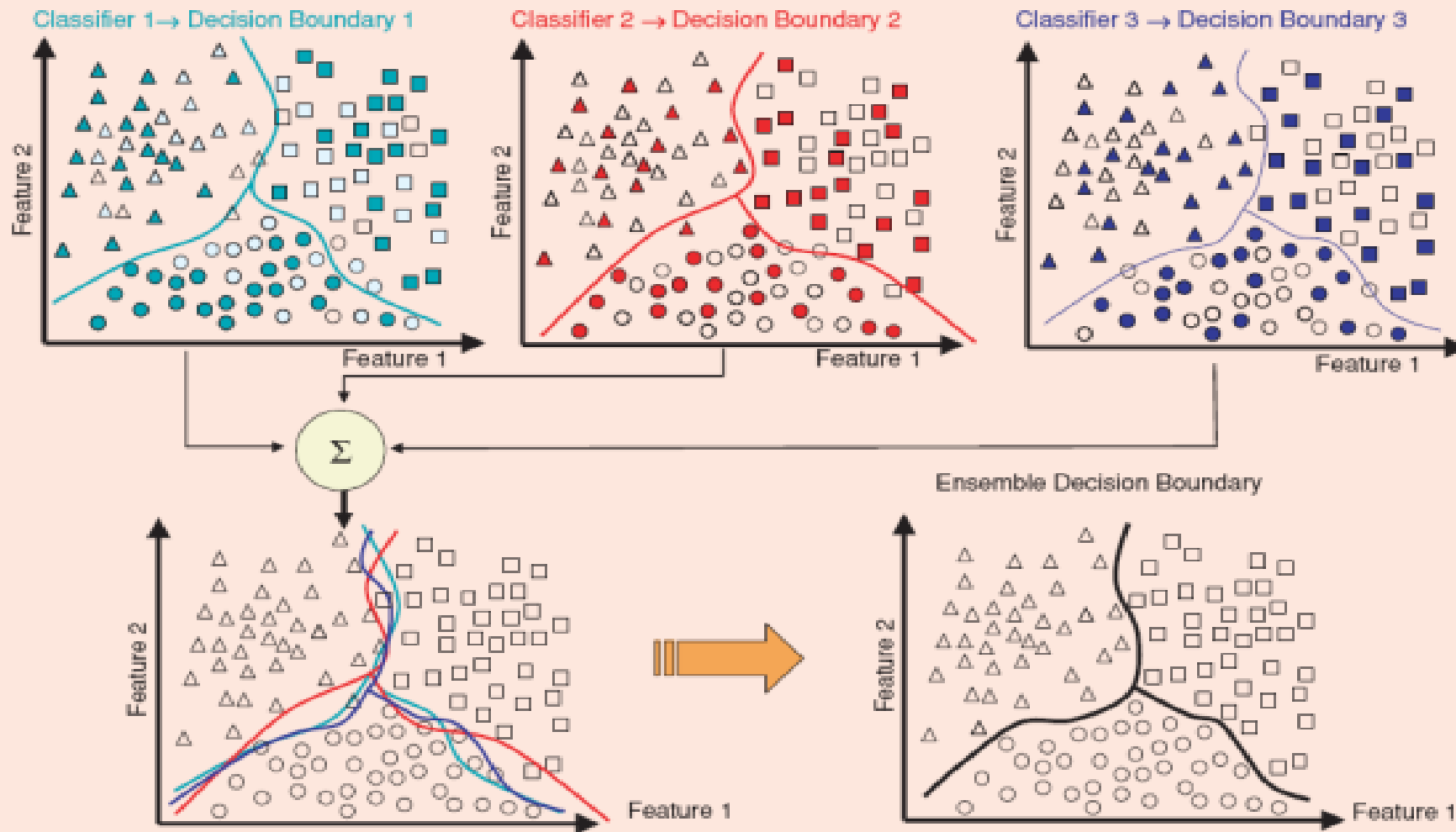
$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} : x_i \in \mathbb{R}^n, y_i = d(x_i)$$

$$model, classifier : f : \mathbb{R}^n \to \{0,1\}$$

$$f|_D \approx d|_D$$

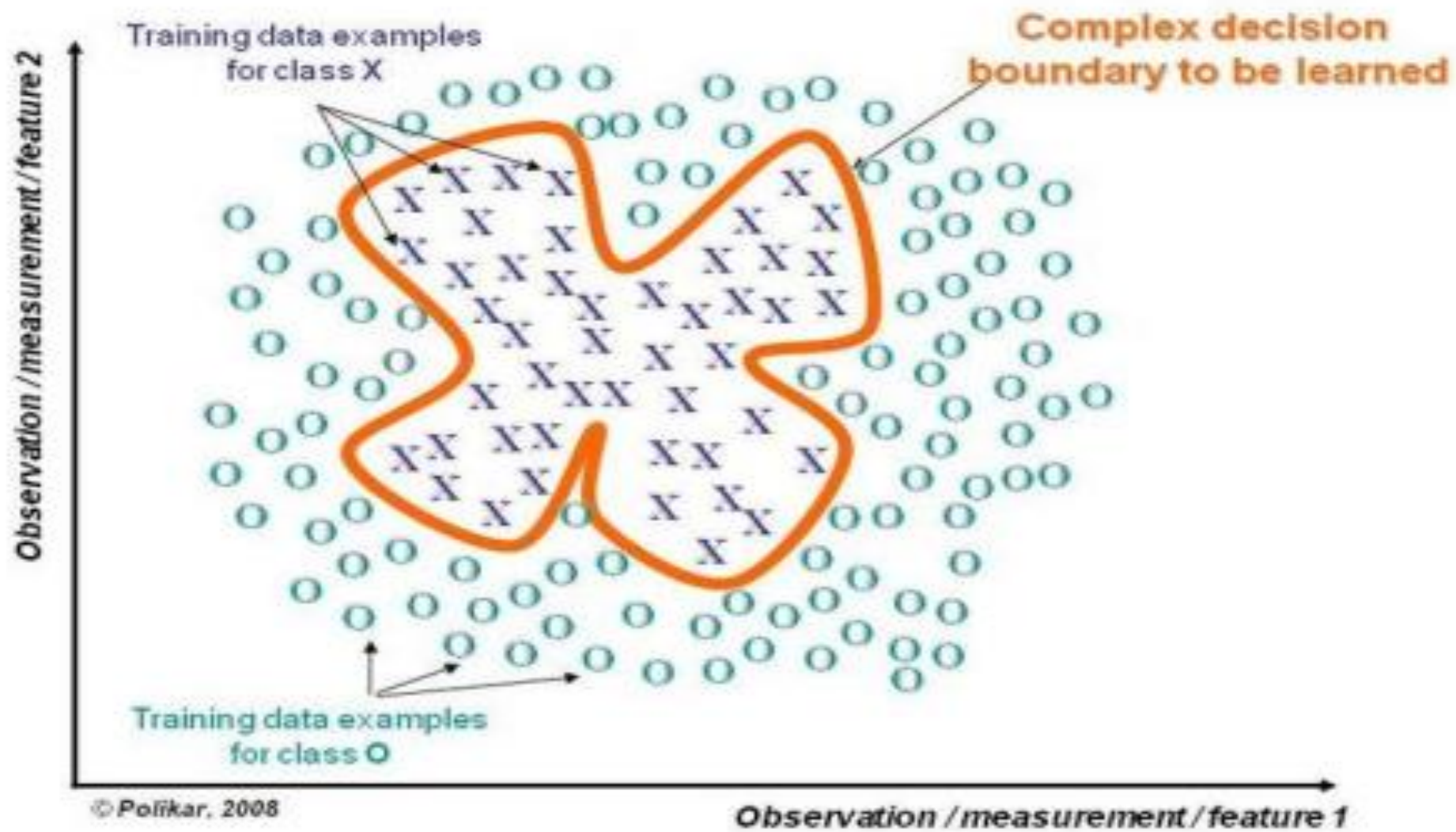$$error_D(f) = \frac{|\{x_i : f(x_i) \neq y_i\}|}{|D|}$$
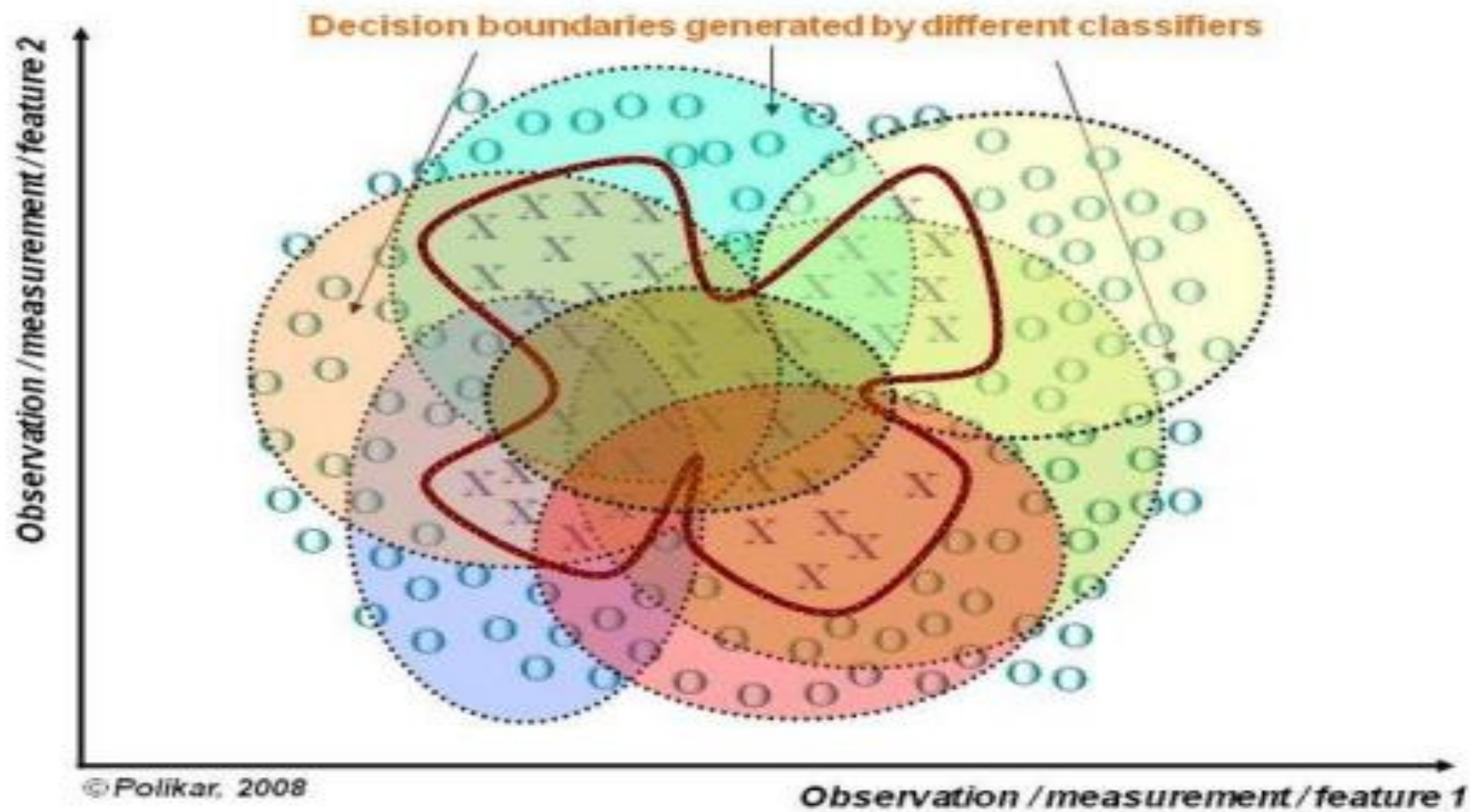
Model Selection

Divide-and-conquer

Divide the problem into a number of subproblems that are smaller instances of the same problem.

Conquer the subproblems by solving them recursively. If they are small enough, solve the subproblems as base cases.

Combine the solutions to the subproblems into the solution for the original problem.

Decision boundaries generated by different classifiers

Observation /measurement/feature 2

Observation /measurement/feature 1

© Polikar, 2008

ENSEMBLE – SUMMARY4