Brief Introduction for Statistical /Machine Learning

黃志勝（Tommy Huang）
義隆電子 人工智慧研發部
國立陽明交通大學 AI學院 合聘助理教授
國立台北科技大學 電資學院合聘助理教授

Introduction for Machine Learning

- What is machine learning?

Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.
[https://en.wikipedia.org/wiki/Machine_learning]

What is the “KEY” in machine learning?

ANS: Data.

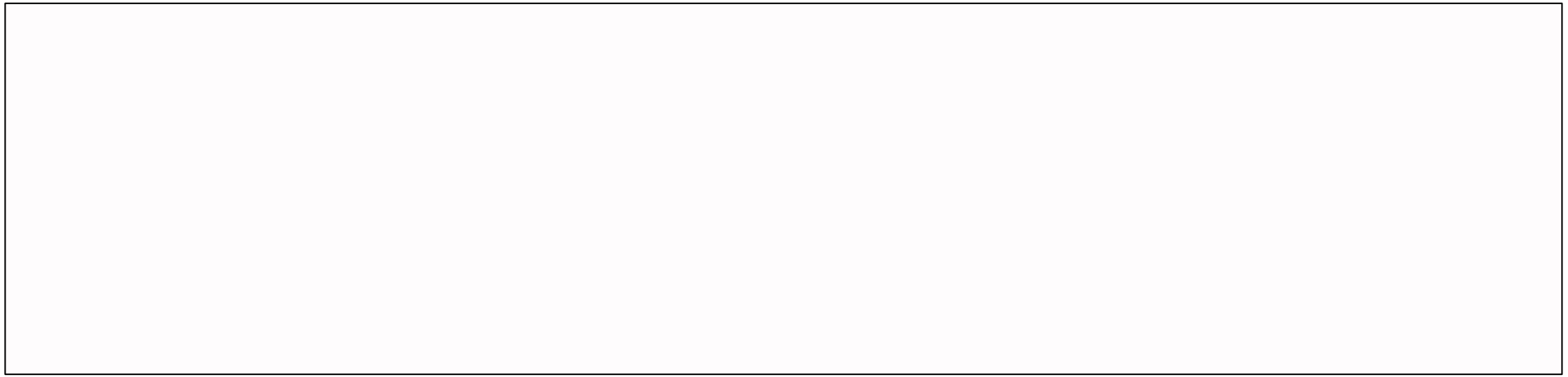
Introduction for Machine Learning

“MOST IMPORTANT” is DATA.

“**Garbage in, garbage out**.”

Introduction for Machine Learning

“MOST IMPORTANT” is DATA.



Data processing is also a science.

Introduction for Machine Learning

Before any projects, please do two right things

1. Specific Problem:
   What problem you want to solve?
2. Right Data (features):
   What is useful data?
   More information more better?

Specific Problem

For instance,

- 1. Spam email detection
- 2. Stock Market Forecast
- 3. Game control
- 4. Heart attack detection

Right Data

For instance,

- Spam email detection: mail content

- Stock Market Forecast: past price,…

- Heart attack detection: ECG signal

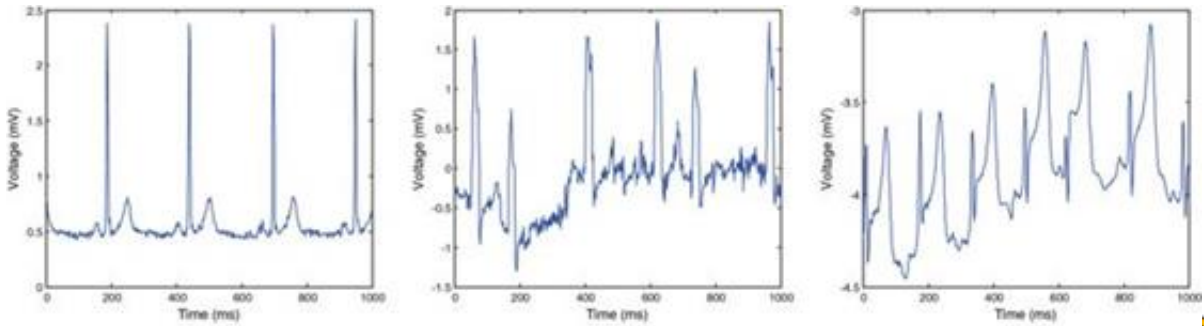
Good day,

I'm Azim Hashim Premji, an Indian business tycoon, investor, and philanthropist. I'm the chairman of Wipro Limited. . I gave away 25 per cent of my personal wealth to charity. And I also pledged to give away the rest of 25% this year new 2019.. I wish to donate $700,000.00USD to every individual. Congratulation, a donation of $700,000.00 has been made to you. If you are interested in my donation, do contact me directly via:
[redacted]@gmail.com for more info. Note: Opportunities comes, but once.

You can also read more about me via the link below

http://en.wikipedia.org/wiki/Azim_Premji

Thank You



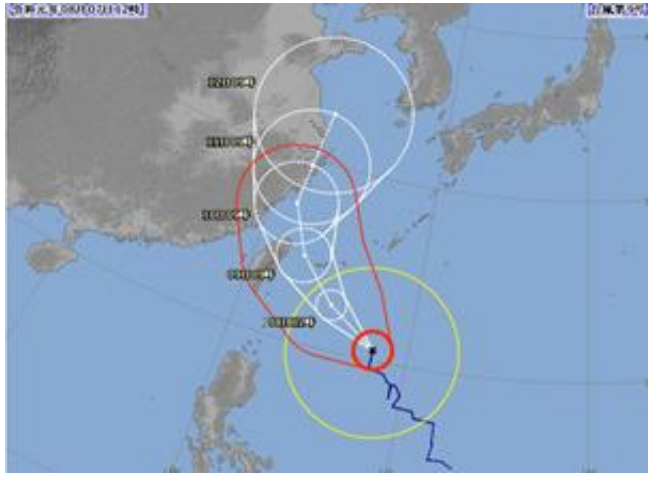
What tasks does ML do?

Depends on what you want it do.

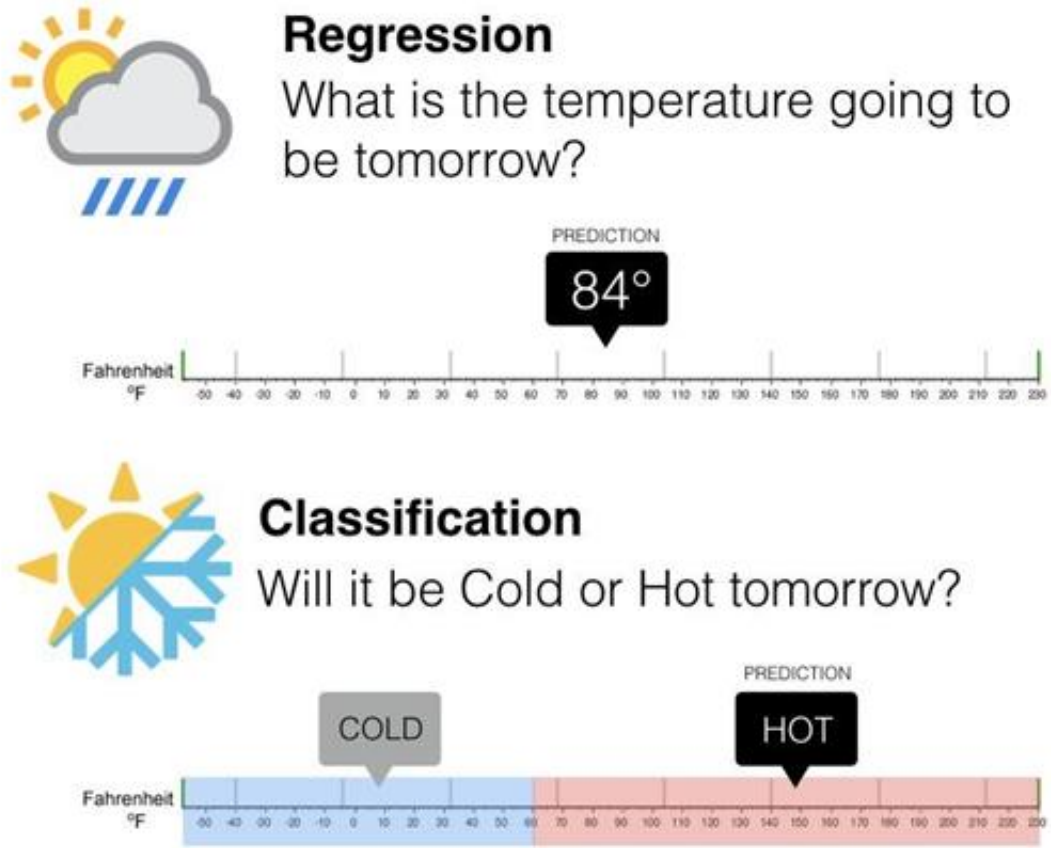
Basically,

Classification

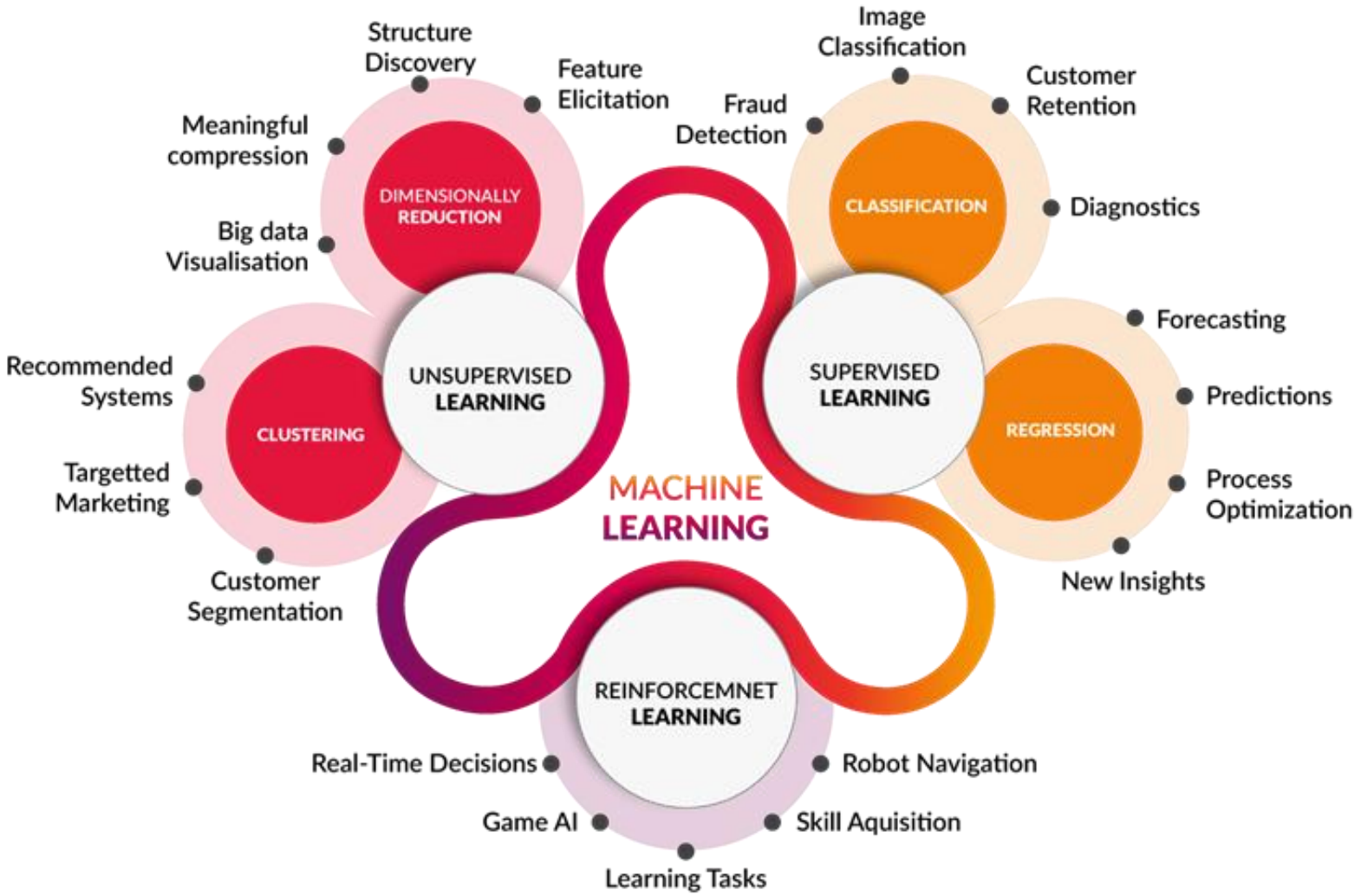
Regression



Classification & Regression



Usually classify the machine learning as …



It's beautiful and almost correct, but NOT really truly classification.

As I know, DAFE (dimension reduction approach) is a supervised learning algorithm.

Scikit-learn classification for ML

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.
Algorithms: SVM, nearest neighbors, random forest, ... — Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.
Algorithms: SVR, ridge regression, Lasso, ... — Examples

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes
Algorithms: k-Means, spectral clustering, mean-shift, ... — Examples

Dimensionality reduction

Reducing the number of random variables to consider.

Applications: Visualization, Increased efficiency
Algorithms: PCA, feature selection, non-negative matrix factorization. — Examples

Model selection

Comparing, validating and choosing parameters and models.

Goal: Improved accuracy via parameter tuning
Modules: grid search, cross validation, metrics. — Examples

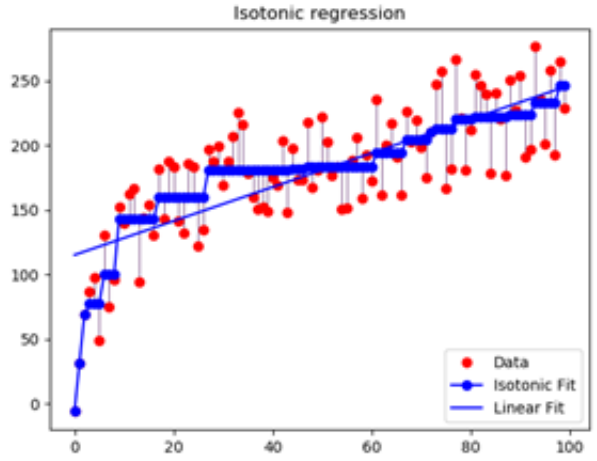
Preprocessing

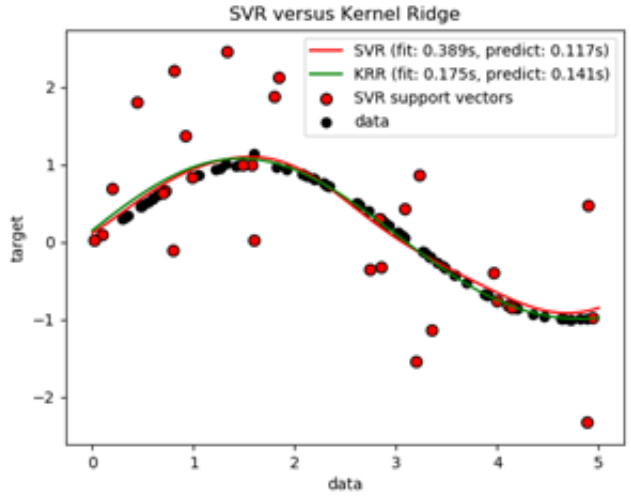
Feature extraction and normalization.

Application: Transforming input data such as text for use with machine learning algorithms.
Modules: preprocessing, feature extraction. — Examples

Regression

- Predicting a continuous-valued attribute associated with an object.
- Linear Regression
- Regularized Regression:
  - Ridge Regression: L2-norm linear regression
  - LASSO (least absolute shrinkage and selection operator): L1-norm linear regression
  - Elastic Net: L1+L2-norm linear regression
- SVR (support vector regression)
- Neural Network

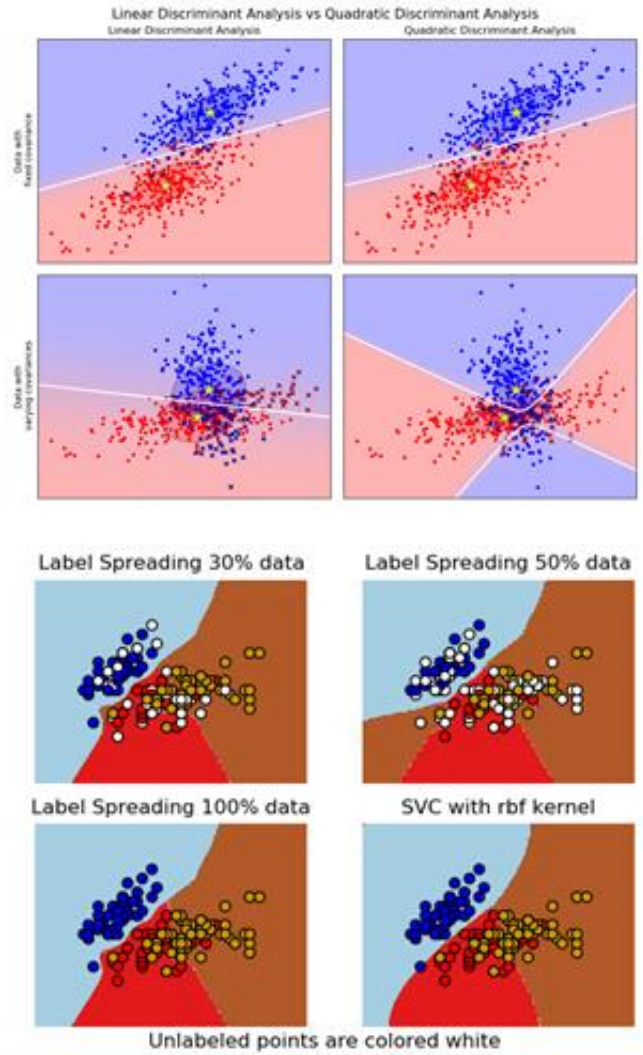




Classification

Identifying to which category an object belongs to.

- Logistic Regression
- Support Vector Machine
- Nearest neighbors
- Random forest
- Neural Network

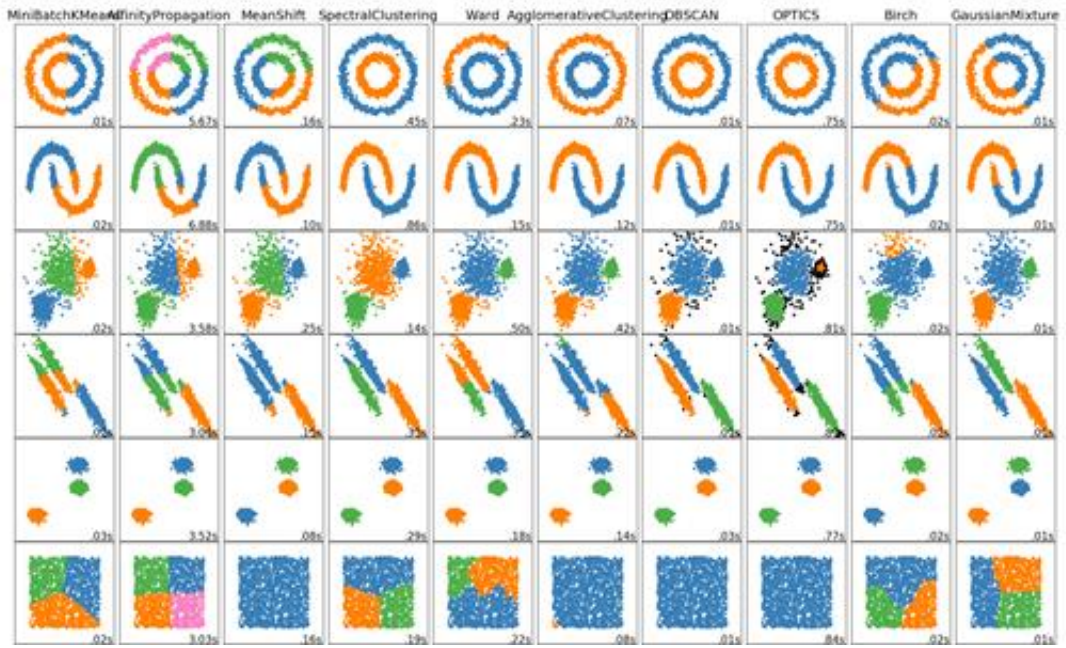


Clustering

Automatic grouping of similar objects into sets.

(物以類聚)

- k-Means
- spectral clustering
- Gaussian mixtures
- Neural Network



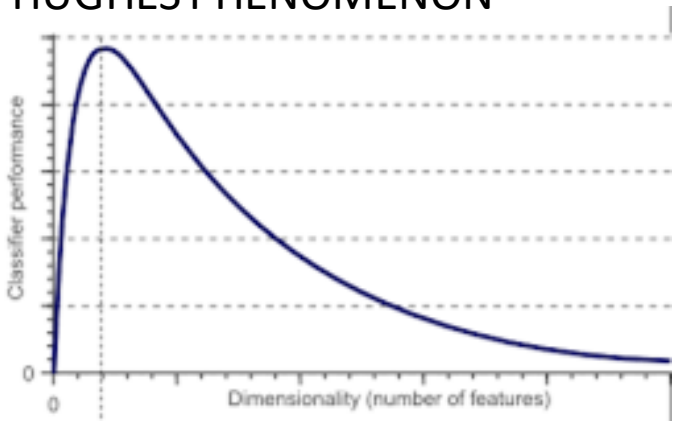
Dimension Reduction

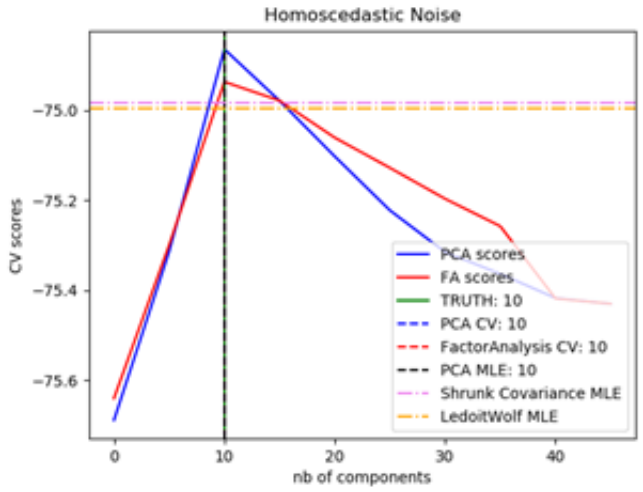
Reducing the number of random variables to consider.

To overcome curse of dimensionality.

- Principal component analysis (PCA)
- Independent component analysis (ICA)
- Canonical component analysis (CCA)
- Non-negative matrix factorization
- Discriminant Analysis Feature Extraction(DAFE)
- Neural Network

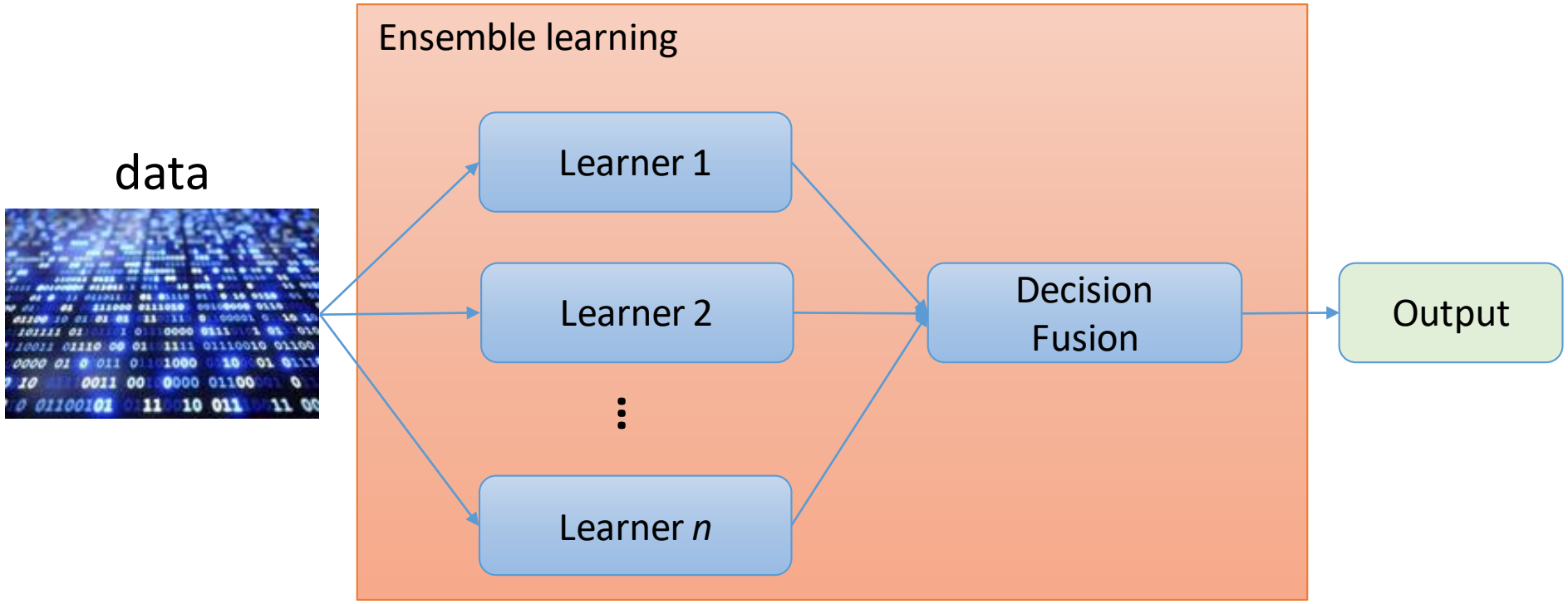
HUGHES PHENOMENON





Ensemble learning

Multiclassification System (三個臭皮匠勝過一個諸葛亮)



Introduction for these topic

1. Regression
2. Classification (Linear and Quadratic Discriminant Analysis)
3. Dimension Reduction (PCA)
4. Back Propagation Network (BPN)