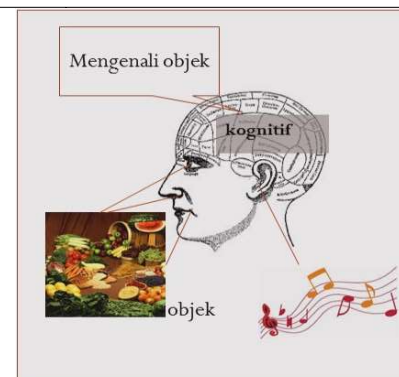
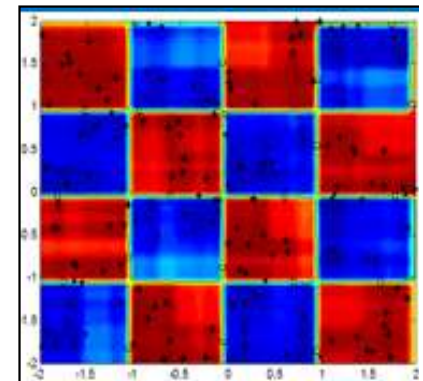




Machine Learning dan Aplikasinya di Biomedis

Elly Matul Imah





ELLY MATUL IMAH
Universitas Negeri Surabaya
Matematika

Subjects/Areas:

ID

Machine Learning Computational Intelligence Pattern Recognition
Biomedical Engineering

1982 Lahir
Bojonegoro, 5 April 1982

2004 Pendidikan
S1 Matematika,
Institut Teknologi Sepuluh November (ITS)
Surabaya

2009 S2 Computer Science
Universitas Indonesia

2014 S3 Computer Science
Universitas Indonesia

- Ketua Program Studi S1 Matematika FMIPA Unesa (2016-2019)
- Sekretaris Pusat HKI dan Publikasi, LPPM Unesa (2019-skrng)
- Tim Strategis Wakil Rektor Bidang Akademik Unesa (2019-skrng)
- Asesor Akreditasi Jurnal Nasional Kemenristekdikti (2018-skrng)
- Reviewer Jurnal Internasional (2014-skrng): IEEE Transaction Neural Network, IEEE SMC, IET Image Processing, IET Biomedical System

Outline

1. Machine Learning
2. Aplikasi Machine Learning di Biomedis:
 - a. Deteksi Jantung
 - b. Neurocognition
3. Prospek dan open problem in machine learning

Apakah semua
object pada
Gambar 1
adalah pohon?



AMERICAN ELM



GINKGO



WEeping WILLOW



SPRUCE



LARCH



CANYON LIVE OAK



TELEPHONE POLE



BIRCH



MONTEREY CYPRESS



SCRUB PINE



DATE PALM



RED MANGROVE

Gambar 1

Machine Learning?

- Salah satu cabang Kecerdasan Buatan yang fokus pada merancang dan mengembangkan algoritma yang membuat komputer/ mesin mampu berevolusi berdasarkan perilaku data, **atau** bisa juga dikatakan membuat komputer mampu menemukan pola dengan menggali informasi dari data yang diberikan secara otomatis.

Kenapa diberi istilah/ nama Machine Learning?



- Learning -> komputer diberikan inisialisasi awal suatu model pengenalan pola dan beberapa data kemudian menemukan cara bagaimana membuat model menjadi lebih baik.
- Machine -> karena komputer belajar secara otomatis, tanpa intervensi dari manusia kecuali inisialisasi awal model dan menyediakan data

Pattern Recognition?

“Sesungguhnya power dari pemikiran manusia didasarkan pada kemampuan untuk mengenali pola. Komputer masa depan dirancang untuk memiliki kemampuan mengenali pola sehingga menyerupai kemampuan manusia.”

(Ray Kurzweil, NY Times, Nov 24, 2003).

Ex: Manusia membangun pola dari pengalaman *“kalau kita memakan jamur Amanita atau death cap, maka kita akan keracunan”*

“Permasalahan mencari pola data adalah suatu hal yang fundamental dan telah menjadi ‘history’ yang berhasil.”
Bishop

Konsep Pengenalan Pola

- **Pengenalan pola** yang dilakukan manusia merupakan proses **pembentukan persepsi** dari suatu **objek** atau kejadian kemudian **mengolahnya** menjadi sesuatu yang **memiliki arti** secara keseluruhan.
- Kemampuan manusia untuk mengenali pola bergantung pada naluri dan sistem kognitif yang dimiliki.
- Dalam **Ilmu Komputer** khususnya bidang **Machine Learning**, pengenalan pola adalah proses pembelajaran pada mesin untuk mengenali atau mengklasifikasikan objek berdasarkan data pelatihan

Apakah Machine learning itu “magic” atau “alchemy”?



Ali Rahimi, di NIPS 2017:
“machine Learning is just like an
alchemy “



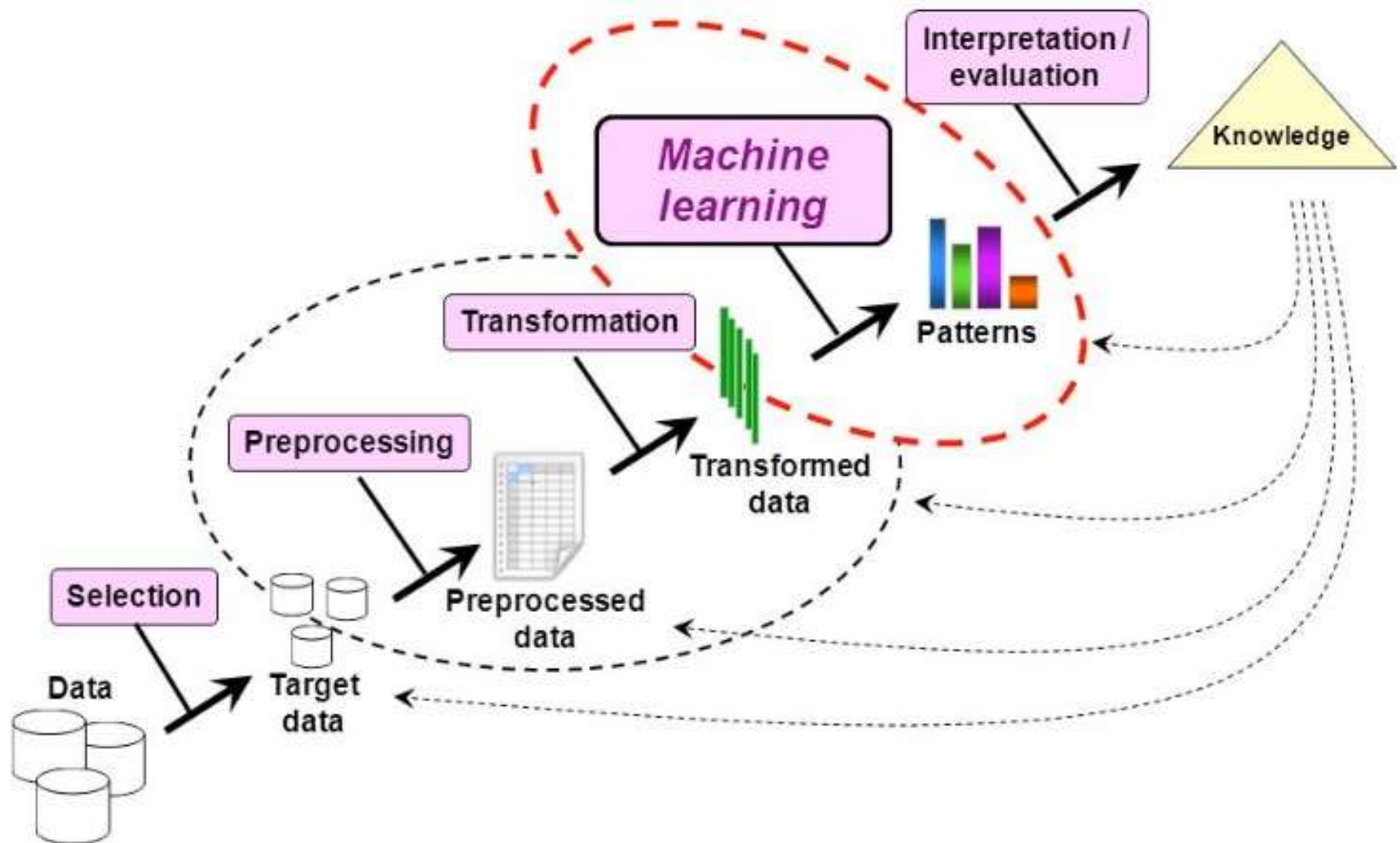
Alchemy “worked”



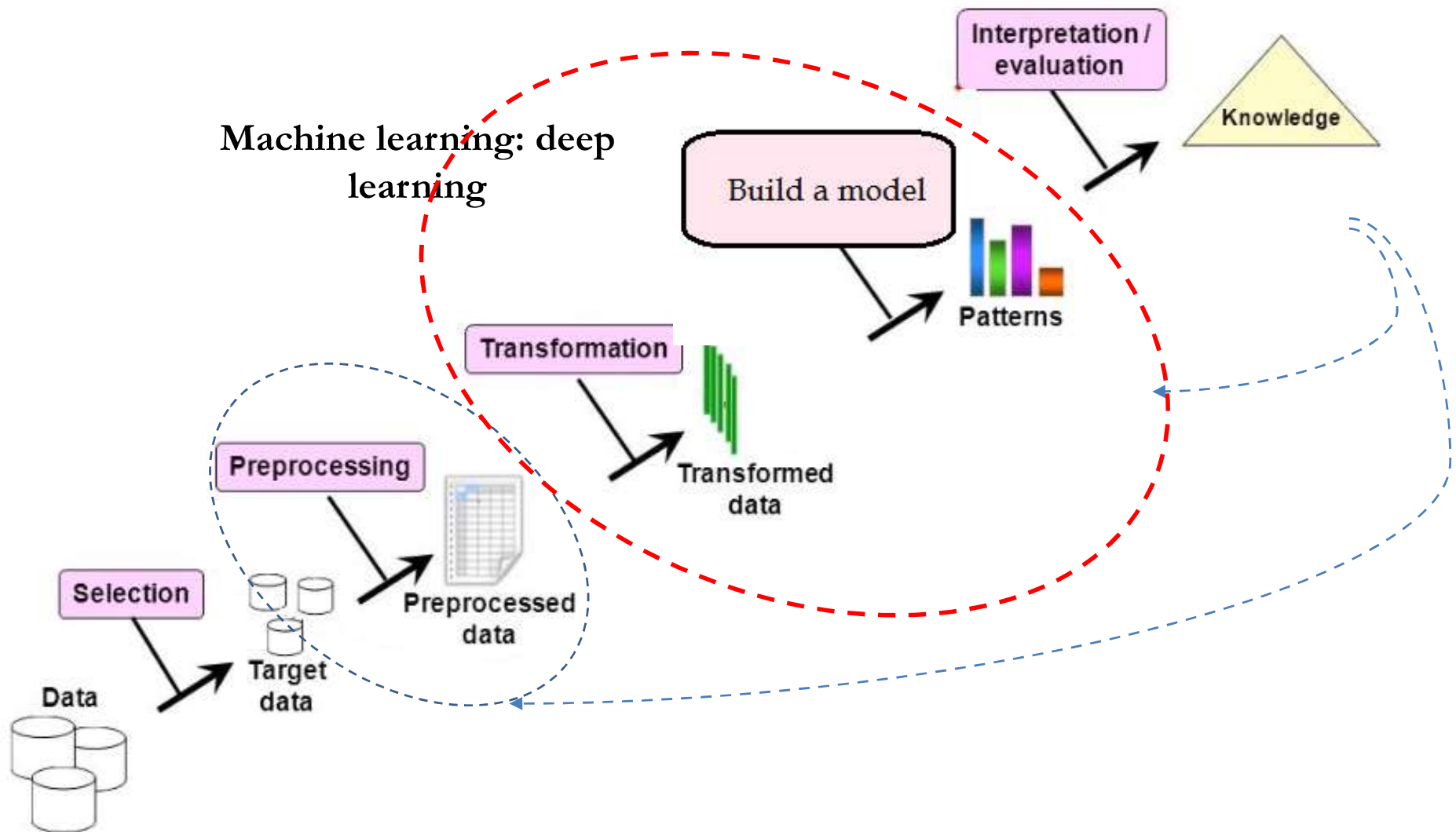
No, more like gardening

- **Seeds** = Algorithms
- **Nutrients** = Data
- **Gardener** = You
- **Plants** = Programs

Tahapan 'mesin' memperoleh Knowledge



Tahapan 'mesin' memperoleh Knowledge

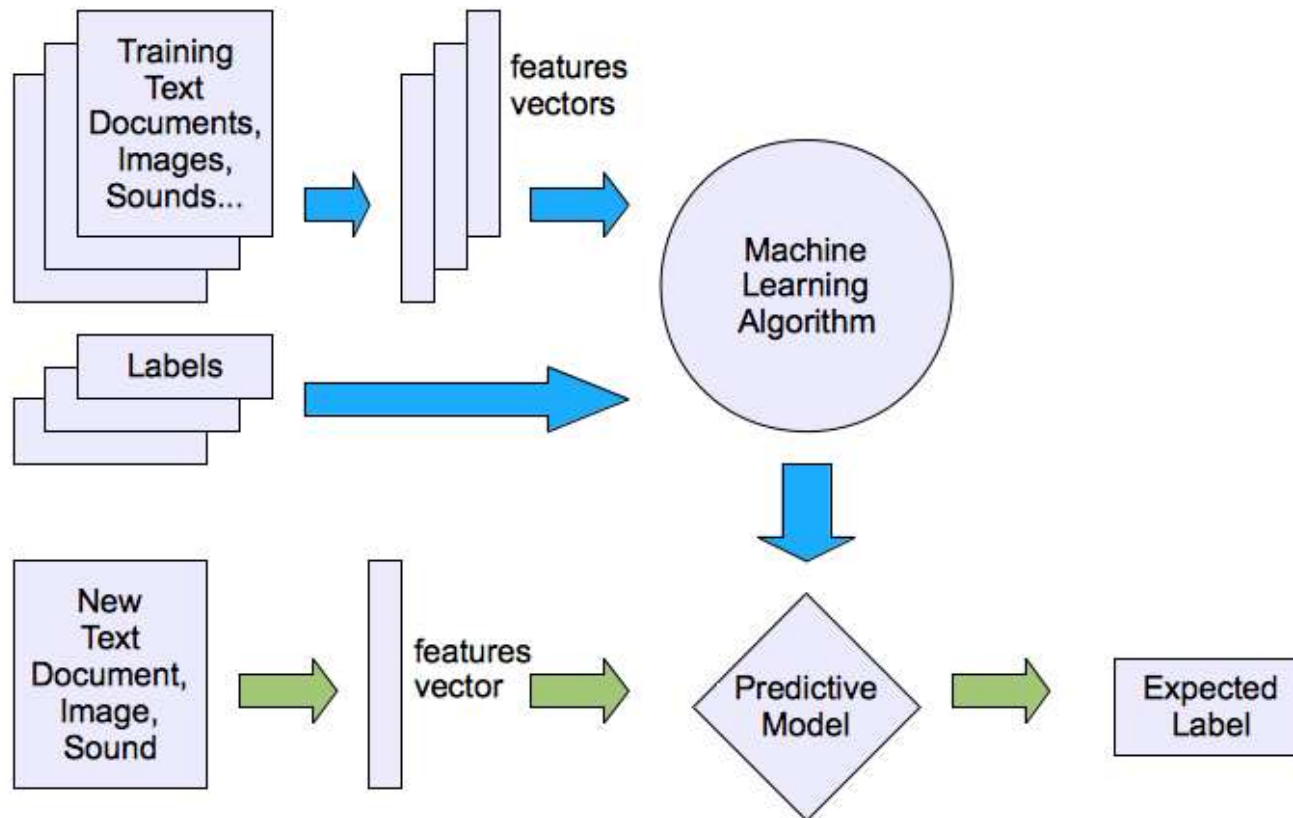


Type Learning pada Machine Learning

- **Supervised (inductive) learning**
 - Training data includes desired outputs
- **Unsupervised learning**
 - Training data does not include desired outputs
- **Semi-supervised learning**
 - Training data includes a few desired outputs
- **Reinforcement learning**
 - Rewards from sequence of actions

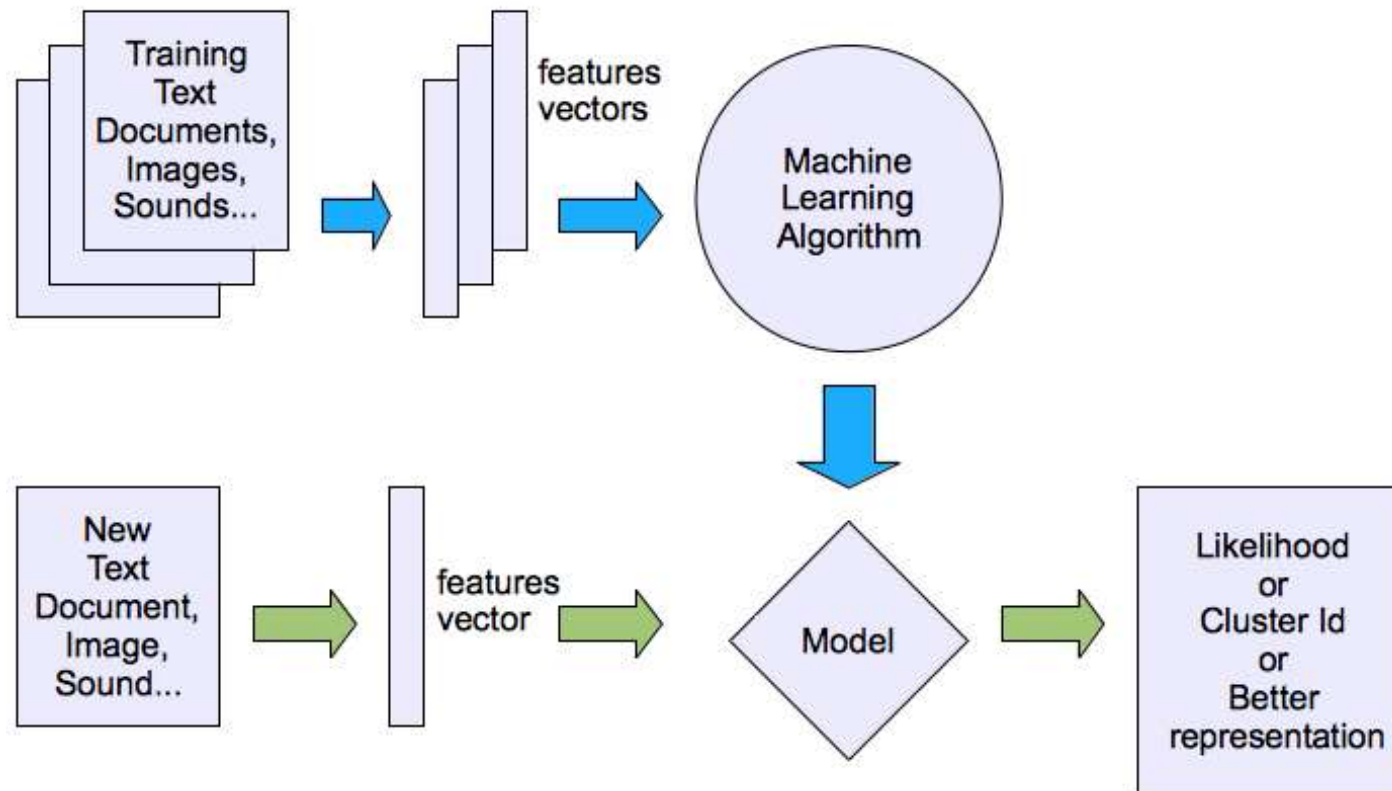
Machine learning structure

- Supervised learning



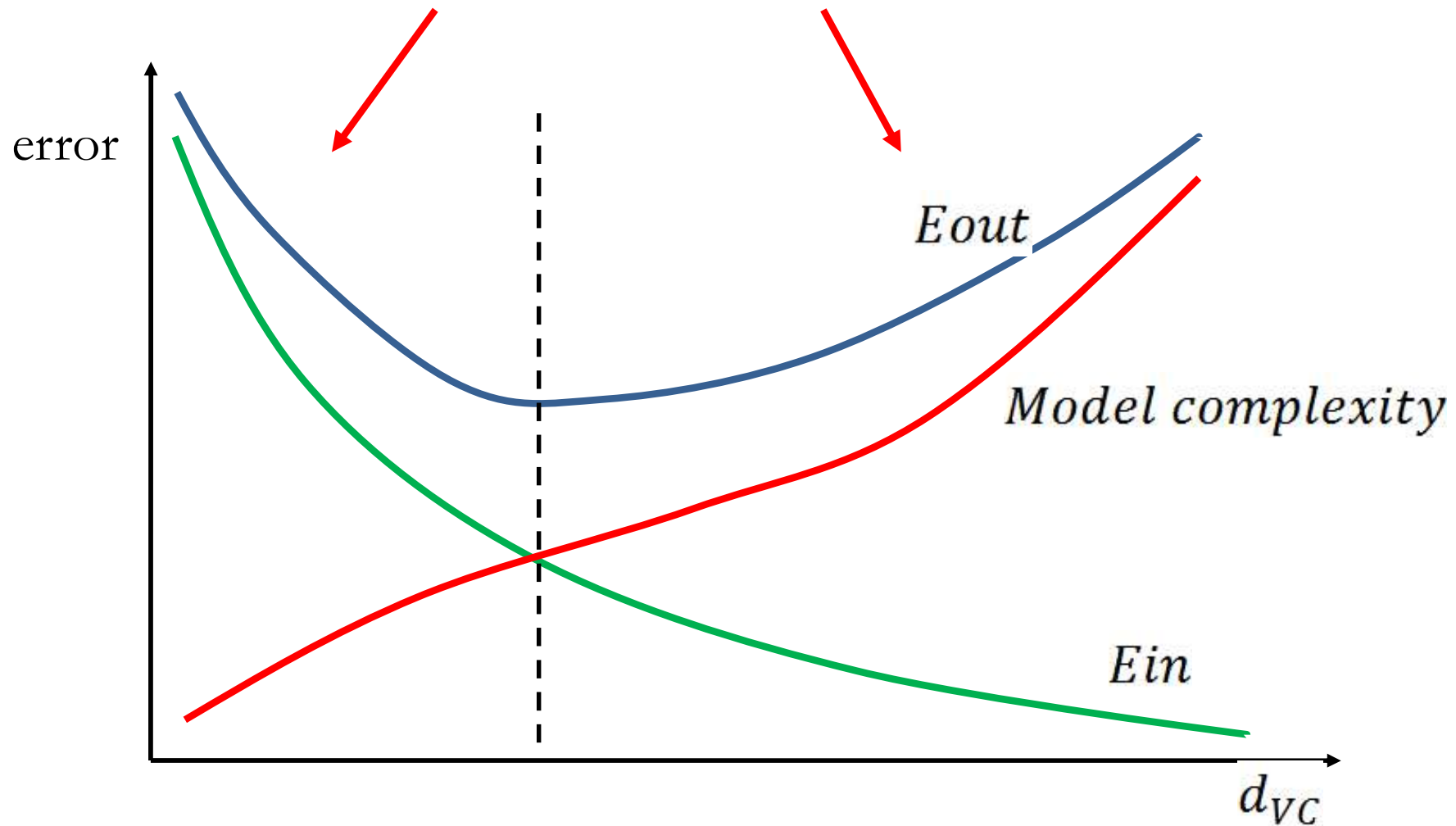
Machine learning structure

- Unsupervised learning



Fenomena dalam machine learning: apa yang dicari?

Under-fitting VS. Over-fitting (fixed N)



Learning techniques

Supervised Learning:

- Support vector machine (SVM)
- Multi-layer perceptron (MLP)
- Extreme Learning Machine (ELM)
- Backpropagation
- Random Forest
- Naïve Bayes
- ...

Unsupervised Learning:

- Fully unsupervised K-means
- Kernel K-means
- SOM
- Hierarchical clustering
- ...

Learning techniques

Semi unsupervised Learning:

- Simple K-means
- SOM with known number of cluster
- ...

Reinforcement Learning:

- LVQ
- GLVQ
- GRLVQ
- ...

Apa bagian terpenting dari Machine learning: “Algorithms”

- Keberhasilan machine learning sangat bergantung pada algoritma yang dipilih.
- Algoritma mengontrol dan menentukan model atau struktur kecerdasan mesin yang dipelajari dari data train.
- Algoritma pembelajaran juga yang menentukan kemampuan mesin mengekstrak informasi dari data train.

Machine Learning
and Pattern
Recognition
(Elly Matul Imah)

Biometrika

- Pengenalan individu berdasarkan detak jantung
- Pengenalan wajah

Biomedika

- Deteksi penyakit jantung Aritmia
- Deteksi endapan trichloroethylene pd hati
- Deteksi epilepsi seized menggunakan data EEG
- **Deteksi Autism Syndrome disorder**
- Deteksi slow learner

Bioinformatika

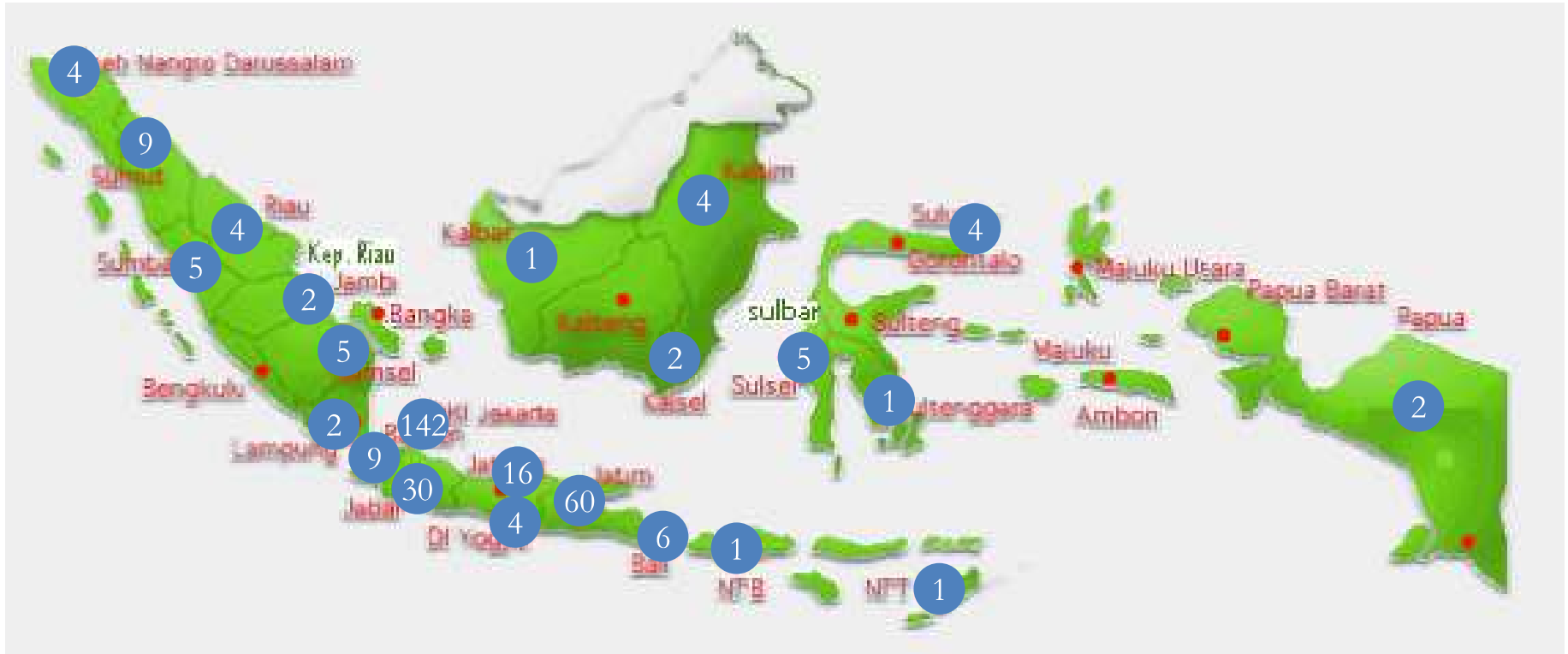
- Drug desain menggunakan machine learning
- BOLD (barcode of life, untuk ikan dan tanaman khas Indonesia)

Social Media

- Hate speech detection

Machine Learning di Biomedis

Cardiologist di Indonesia



source: www.inaheart.org

Based on PERKI databases (*Perhimpunan Dokter Spesialis Kardiovaskular Indonesia*)

- There are approx. 452 cardiologist
- Total Population : 228.5 million (BPS - 2008)

Ratio approx. 1 : 550.000

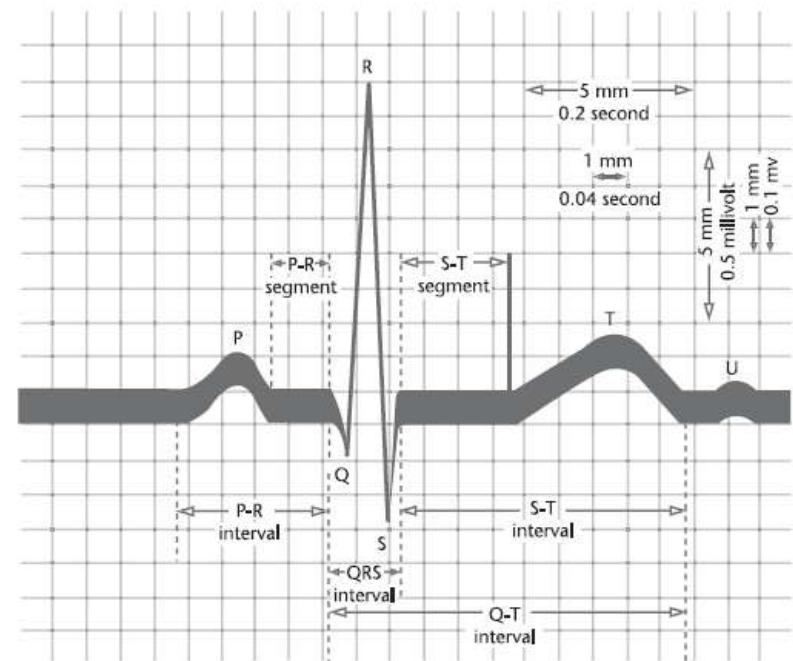
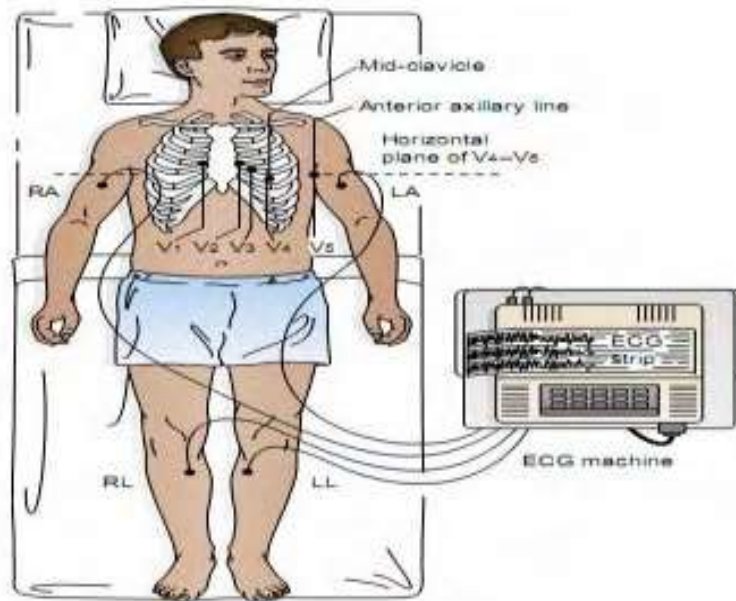
Mengenal Arrhythmia

- **Palpitasi** : perasaan (sensasi) yang tidak menyenangkan yang disebabkan oleh denyut jantung yang tidak teratur.
- Beberapa orang dengan palpitasi-palpitasi (jantung berdebar), tidak menderita penyakit jantung atau kelainan irama jantung (abnormal) dan penyebab jantung berdebarnya tidak

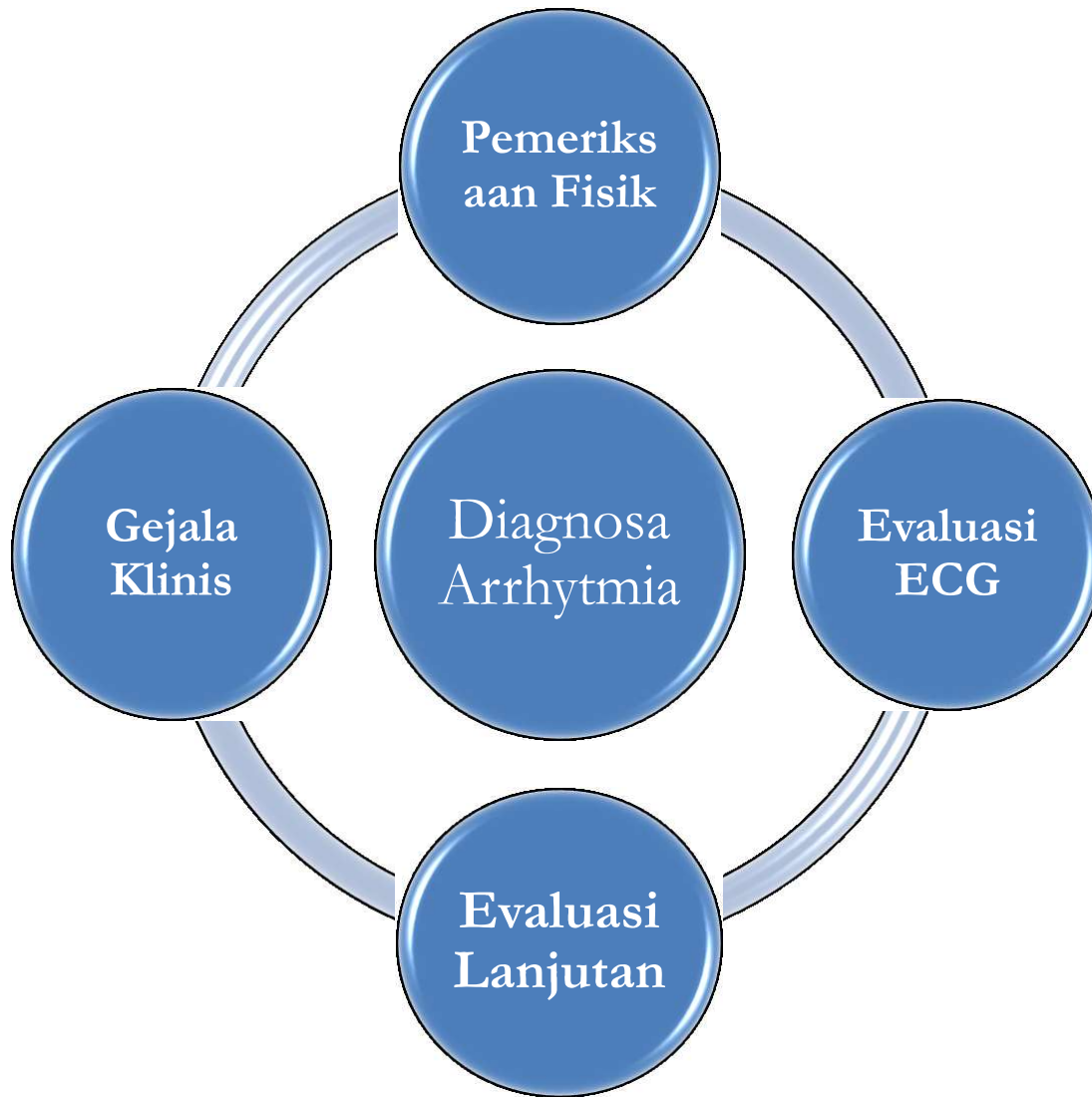


Electrocardiogram (ECG)

- Cara / Alat untuk mendiagnosa penyakit jantung
- Standar 12-lead, namun ada yang 5-lead atau 3-lead



Posisi ECG



Pemeriksaan ECG

“pemeriksaan penunjang dalam menentukan seorang pasien menderita penyakit jantung, khususnya Arrhythmia”

Pendekatan Analisis

- **Analisis Beat**

- Pengenalan Penyakit berdasarkan Beat Tunggal
- Mendeteksi kemunculan gangguan irama jantung mis. RBBB, LBBB, PVC dll

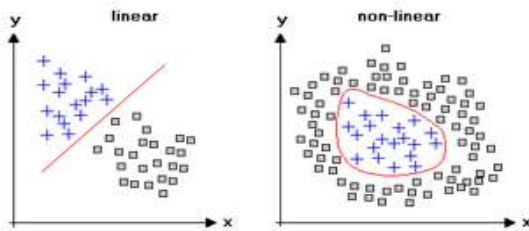
- **Analisis Interval Waktu**

- Peak Interval, mis. RR
 - Mendeteksi gangguan irama jantung seperti tachycardi, bradycardi
- QT Interval
- QRS Interval dll

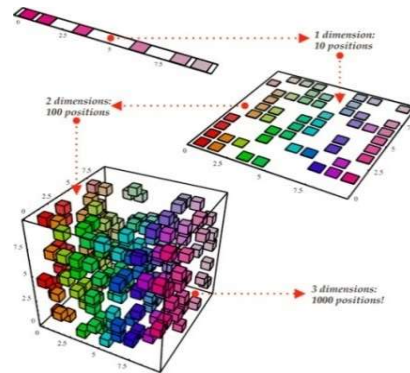
Introduction

Some Challenges in Pattern Recognition:

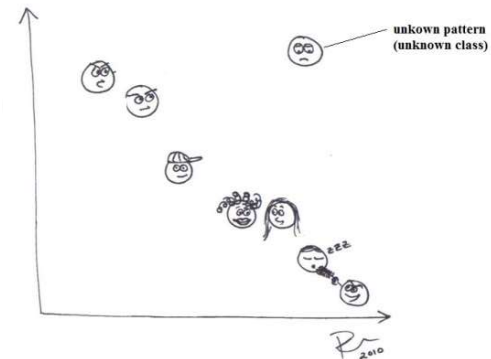
Non Linear Separability



Curse of Dimensionality

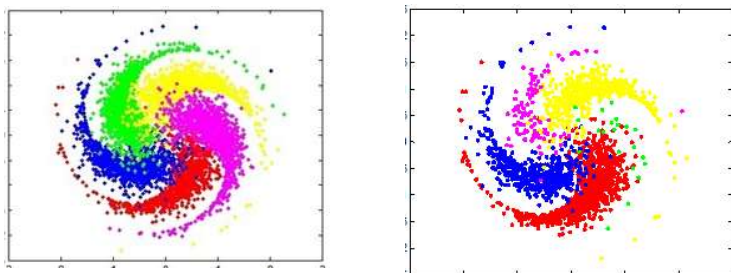


Unknown pattern/ class



<http://moraxella.deviantart.com/art/Outlier-1863>

Imbalanced Data Set



Balanced data set

Imbalanced data set

Too many data with different characteristics in one class



Incremental data or dynamic system



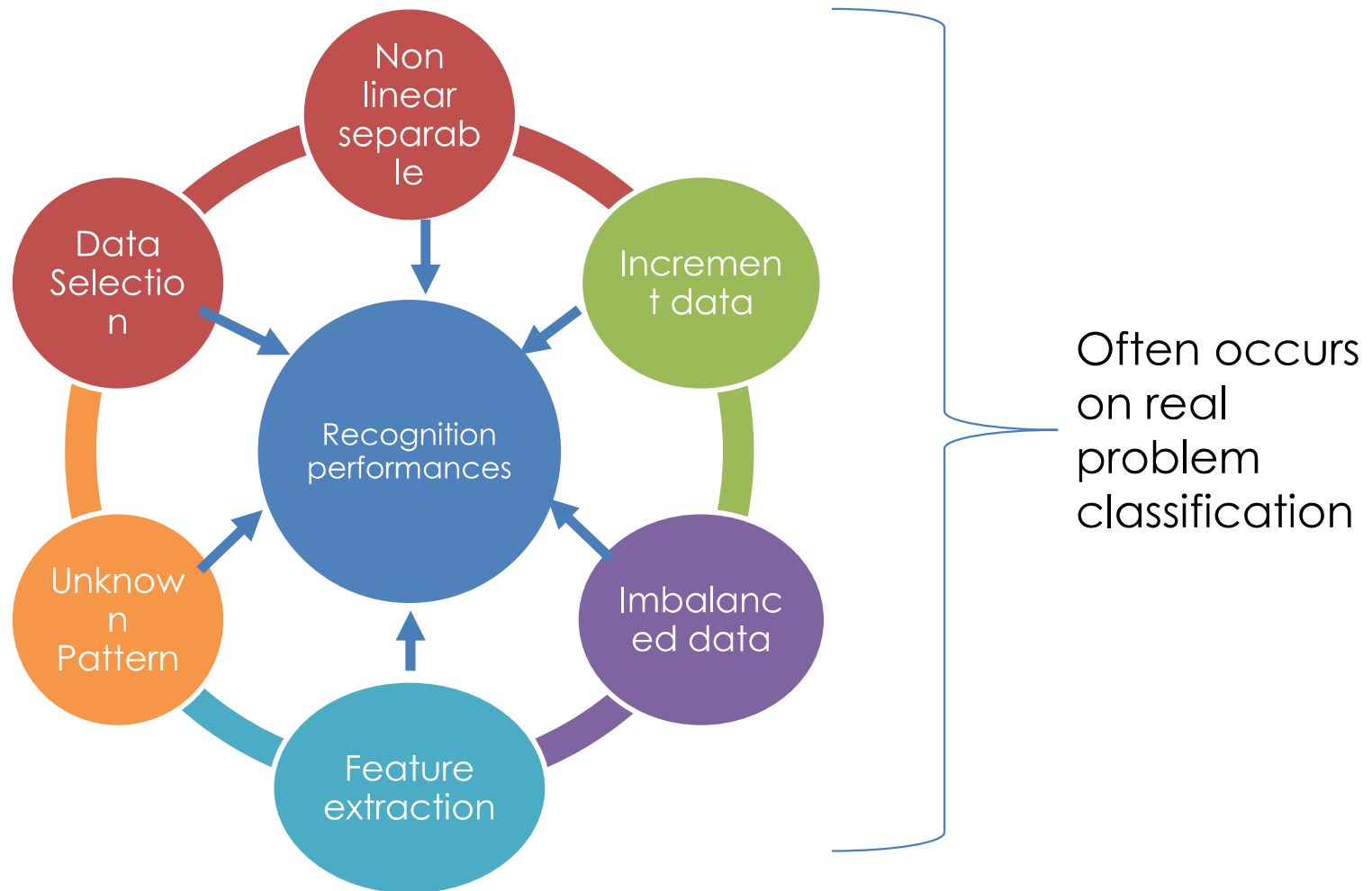
Online learning



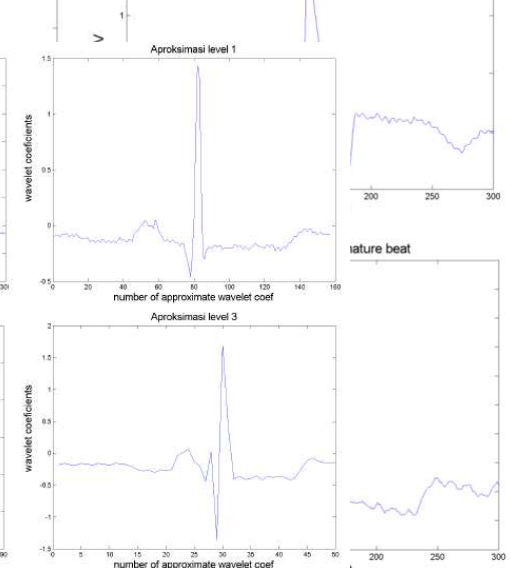
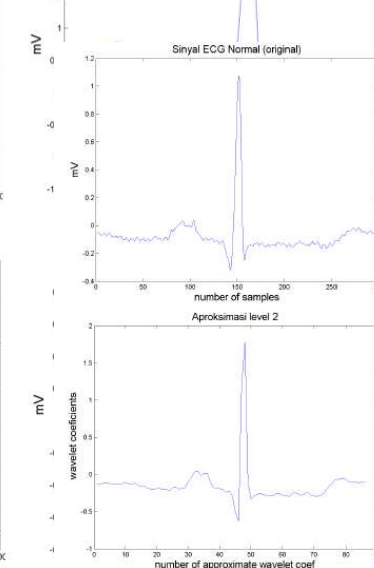
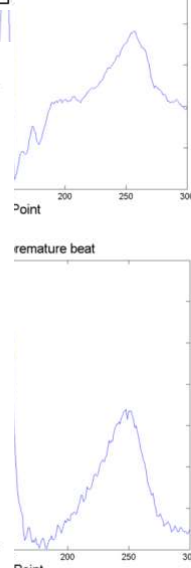
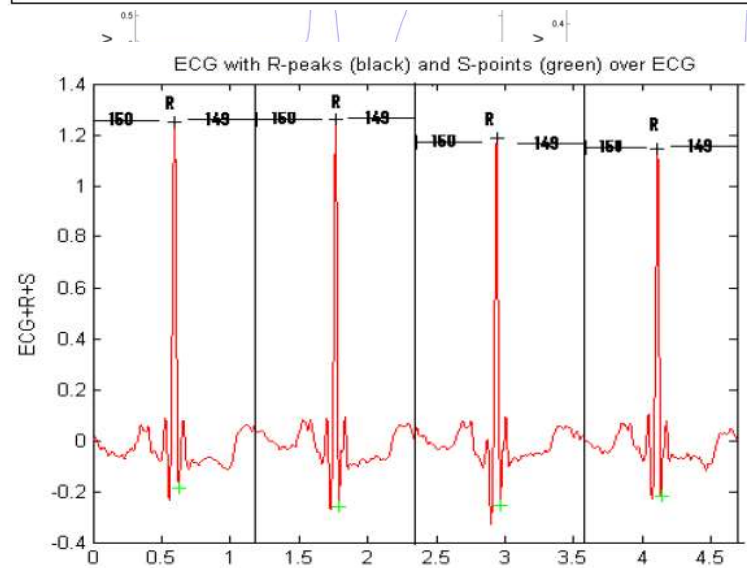
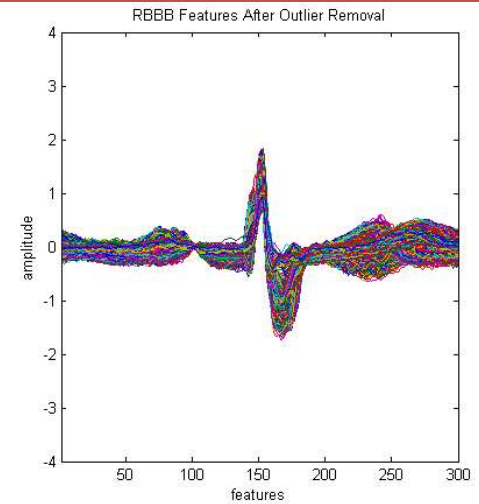
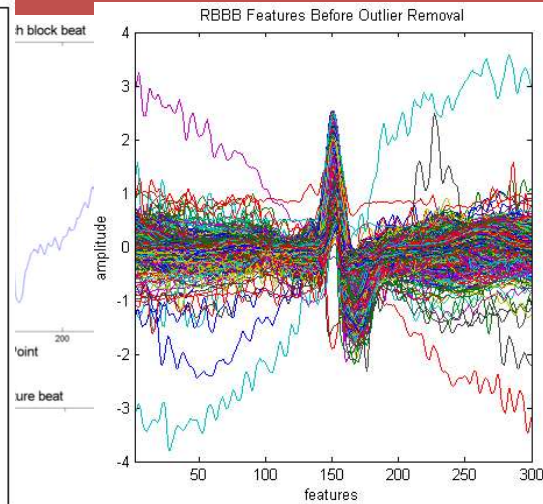
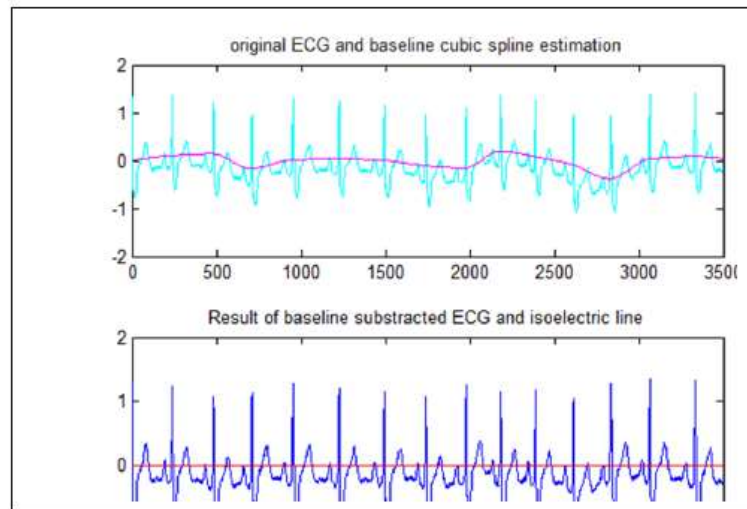
Batch learning

Introduction

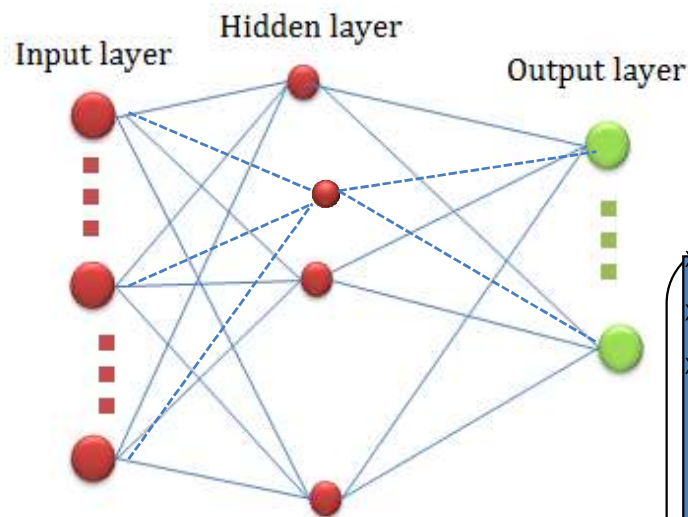
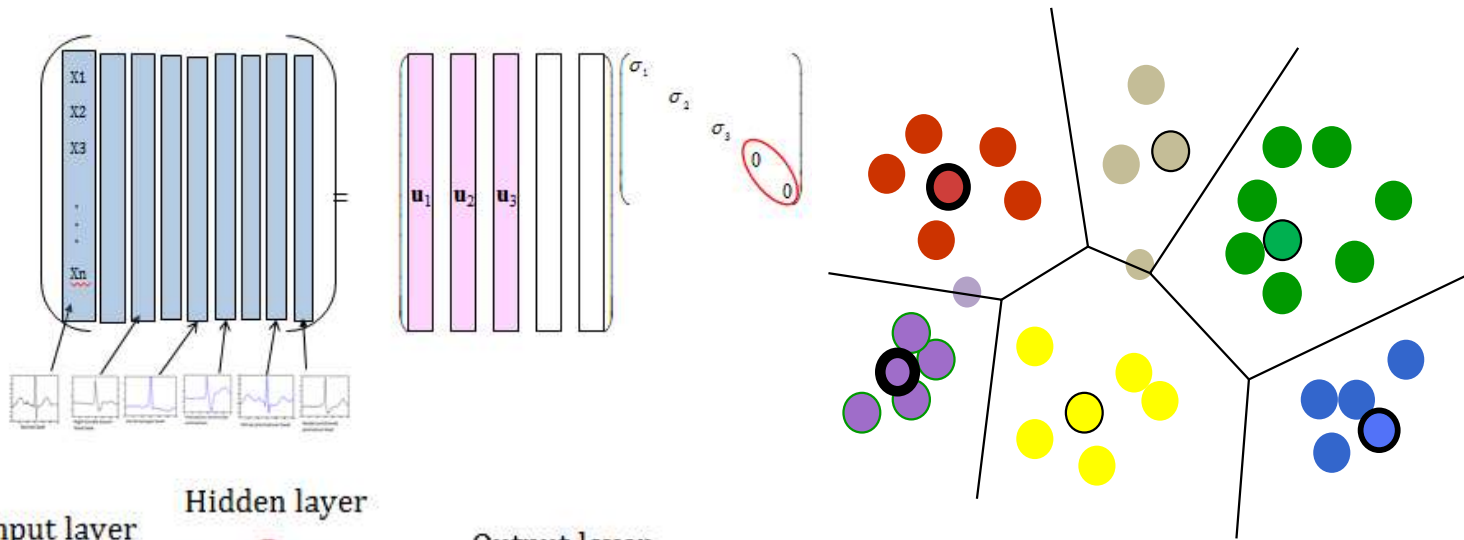
The performance of Pattern Recognition System depends on:



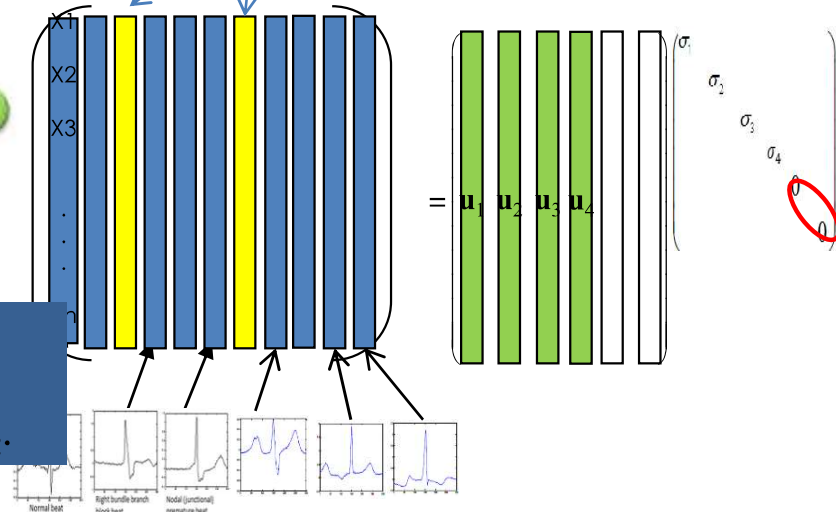
Arrhythmia Recognition System



Adaptive Multilayers Generalized Learning Vector Quantization (AMGLVQ)



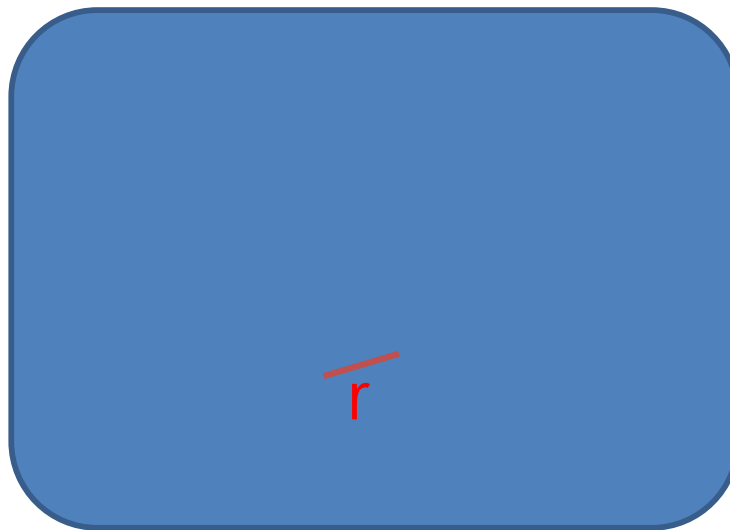
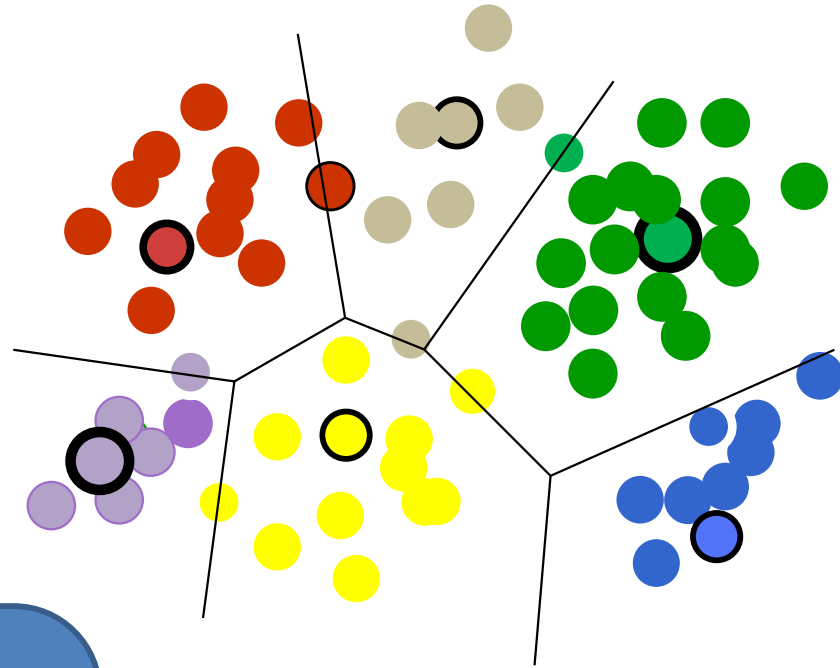
Add new data



MCE relative distance ($\mu(x)$)

$$\mu(x) = \frac{d_1 - d_2}{d_1 + d_2}$$

d_1 and d_2 : distance between x from w_1 and w_2 .



Selected sample
data train

Handling unknown class Data

Find the boundary class

$$r = \frac{\sum_{i=1}^N d_n}{N} + 3\sigma$$

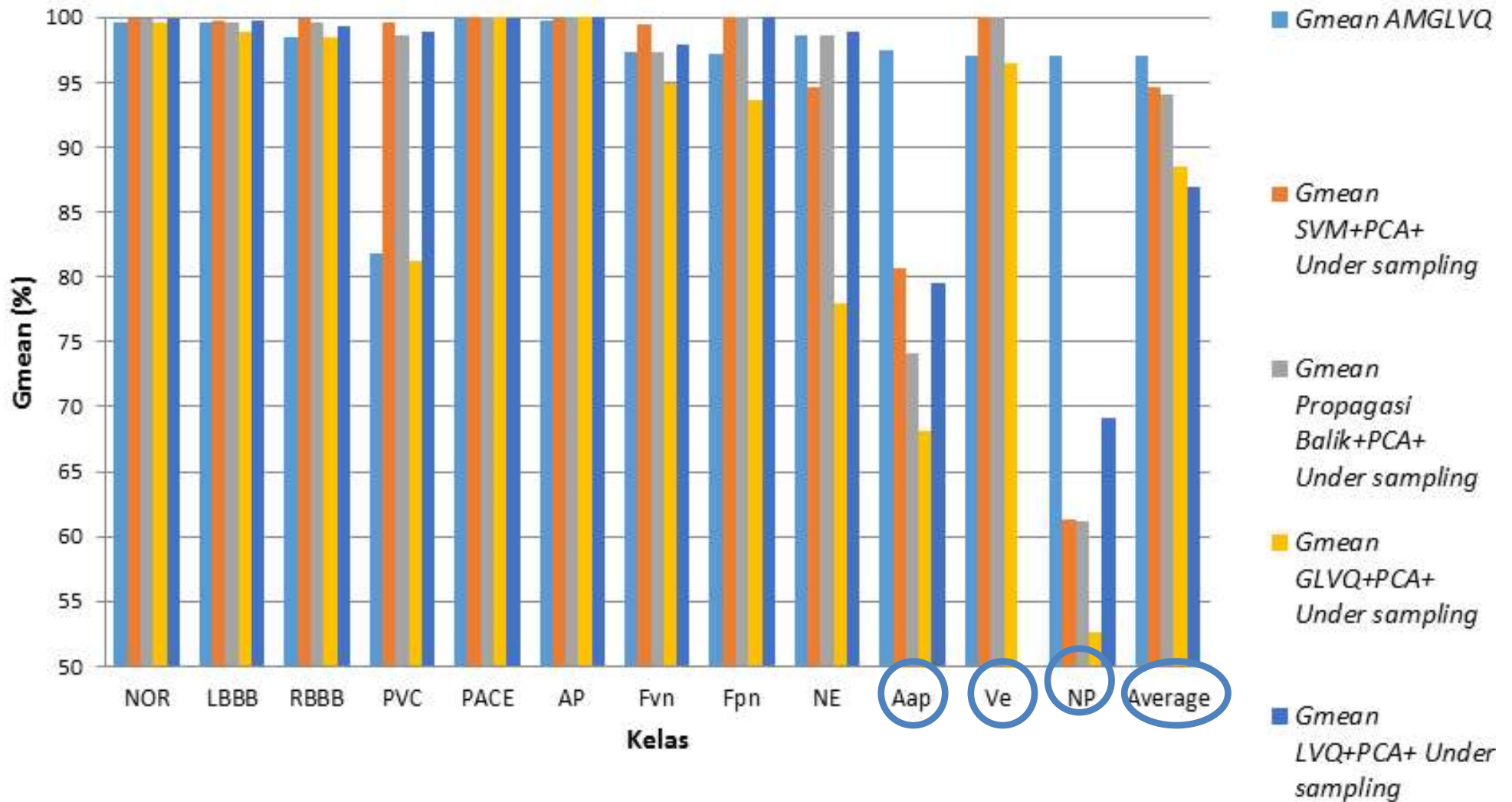
r = threshold

d = distance center class to the edge of class

N = number of data train

σ = variant

Hasil Eksperimen MIT-BIH 12 Kelas

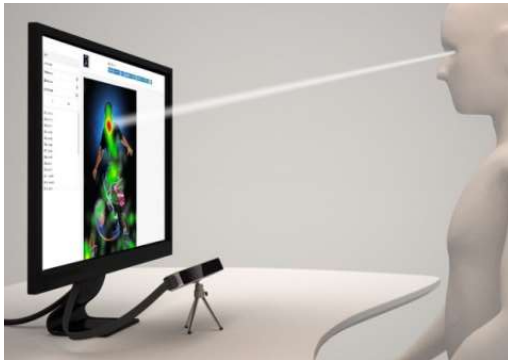


Aplikasi Machine Learning pada Neuro-cognition

Latar Belakang

- Gangguan perkembangan pervasif atau sering dikenal dengan *Pervasive Developmental Disorder* (PDD) merupakan gangguan yang terjadi pada anak berupa kegagalan untuk berkomunikasi, berinteraksi dengan sekitar, dan menunjukkan tanda-tanda perilaku tidak wajar yang dilakukan berulang-ulang dan terus menerus.
- Berdasarkan data *Central of Disease Control* (CDC) Autism dan *Developmental Disabilities Monitoring* (ADM) Network pada tahun 2014 rasio anak penderita autis adalah 1:59, angka ini meningkat dari tahun sebelumnya 1:64.
- Di Indonesia diperkirakan rasio anak penderita autis 3:250 pada tahun 2000 dan jumlah ini meningkat dari tahun ketahun.
- Melihat kondisi tersebut, tentu mengkhawatirkan karena anak-anak ini nanti yang akan menjadi SDM utama diusia produktifnya yang digaungkan menjadi bonus demografi bagi Indonesia untuk mencapai puncak keemasannya.

Permasalahan yang diangkat dalam penelitian



Sistem *eye tracker*
(The Eye Tribe technology,
2016)

**Bagaimana mengembangkan
suatu sistem untuk deteksi dini PDD
menggunakan teknologi
pergerakan mata dan BCI?**

State of the art

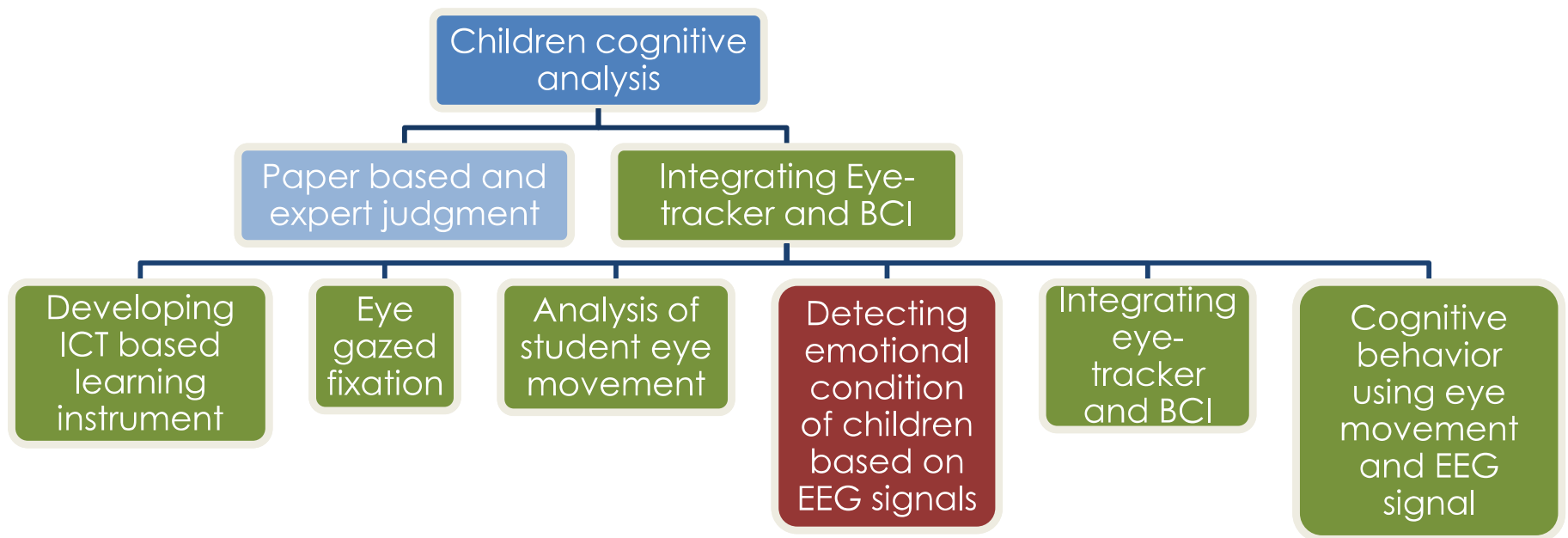
Sinyal EEG untuk klasifikasi *brain control machine* Islam et al., 2016 dan klasifikasi *cognitive state* Zafar et al., 2016.

Salah satu cara klasifikasi emosi manusia adalah dengan sinyal EEG. Penelitian Al-Nafjan et al., 2017 dan Wang et al., 2014.

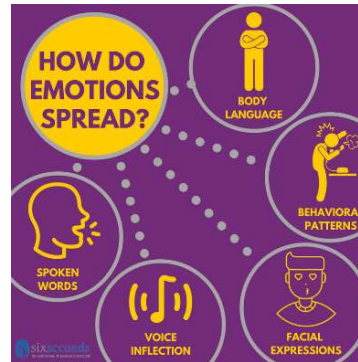
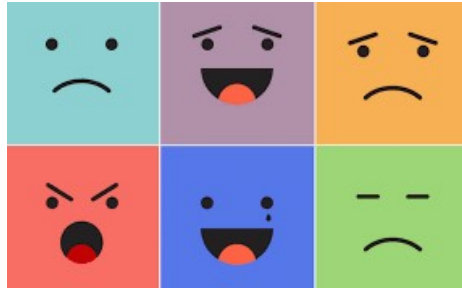
Riby dan Hancock pada tahun 2009 pada penelitiannya menunjukkan bahwa *eye-tracking* dapat digunakan untuk mendeteksi gangguan Sindrom Williams

Forssman dan timnya (2015) melakukan penelitian untuk mengamati fungsi kognitif pada bayi menggunakan eye tracker dari keluarga menengah ke bawah di Finish dan Malawi (Forssman et al., 2015), penelitian tersebut menunjukkan bahwa *eye-tracking* merupakan metode yang sangat bagus untuk mengukur secara dini kemampuan kognitif seseorang walaupun dengan keterbatasan *resource*

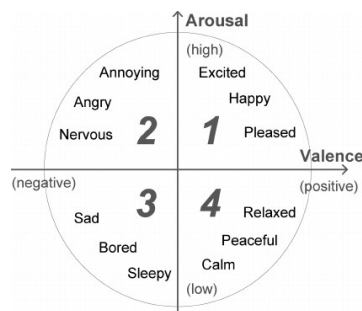
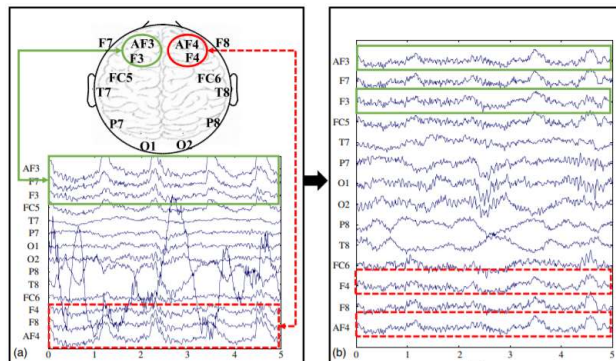
Carrick C. Williams (2015) bersama timya melakukan penelitian untuk melihat hubungan antara pergerakan mata dengan psikologi kognitif, khususnya mengkaji tentang bagaimana pergerakan mata bisa digunakan untuk melihat gangguan mental seseorang



Introduction



The emotional state recognition using facial expressions, voice inflection, spoken words, and body language is not precisely for non-expressive people.



Identifying humans emotional state using electroencephalogram (EEG) signal more precisely than using non-verbal and verbal signals, because emotions are psychological and physiological processes that are connected with personality, motivation, mood, and temperament. The EEG become more precisely for Emotional state classification because humans can manage their expressions, but it is very difficult or impossible to manage human's natural body reactions such as temperature changes, heart rate, or physiological signals

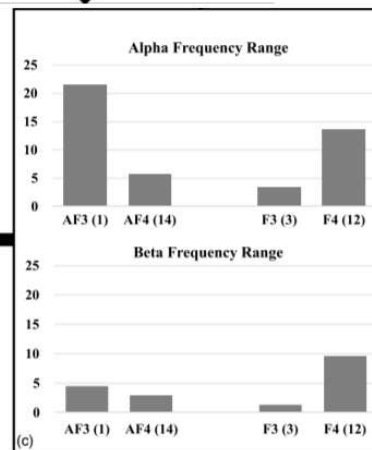
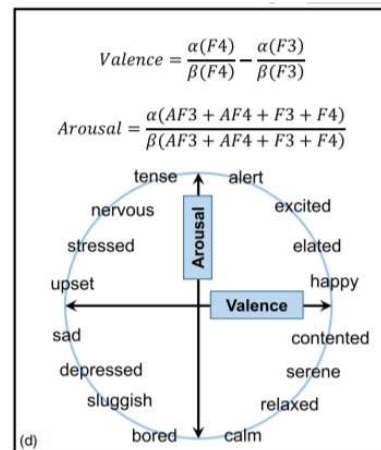
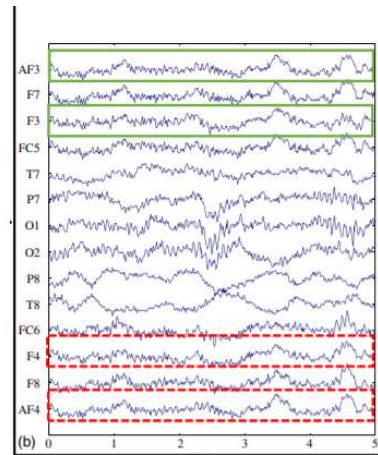
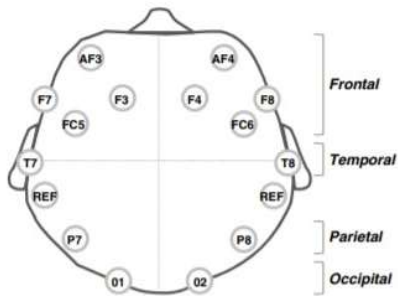
Image sources:

<https://www.6seconds.org/2018/05/29/5-surprising-ways-emotions-spread/>
<https://psychology-spot.com/emotional-allergy/>

Hwang, S., Jebelli, H., Choi, B., Choi, M., & Lee, S. (2018). Measuring Workers' Emotional State during Construction Tasks Using Wearable EEG. *Journal of Construction Engineering and Management*, 144(7), 04018050. doi:10.1061/(asce)co.1943-7862.0001506

Electroencephalogram (EEG) and Emotional State

EEG is a brain activity record in the form of brain waves recorded through the scalp. In the brain, there are millions of neurons that produce a fine electrical voltage, and this voltage makes the electricity in electrodes can detect and receive EEG signals



Feature Extraction

Daubechies Wavelet Transformation

We use discrete wavelet transform to extract the feature contain in the individual signal beat

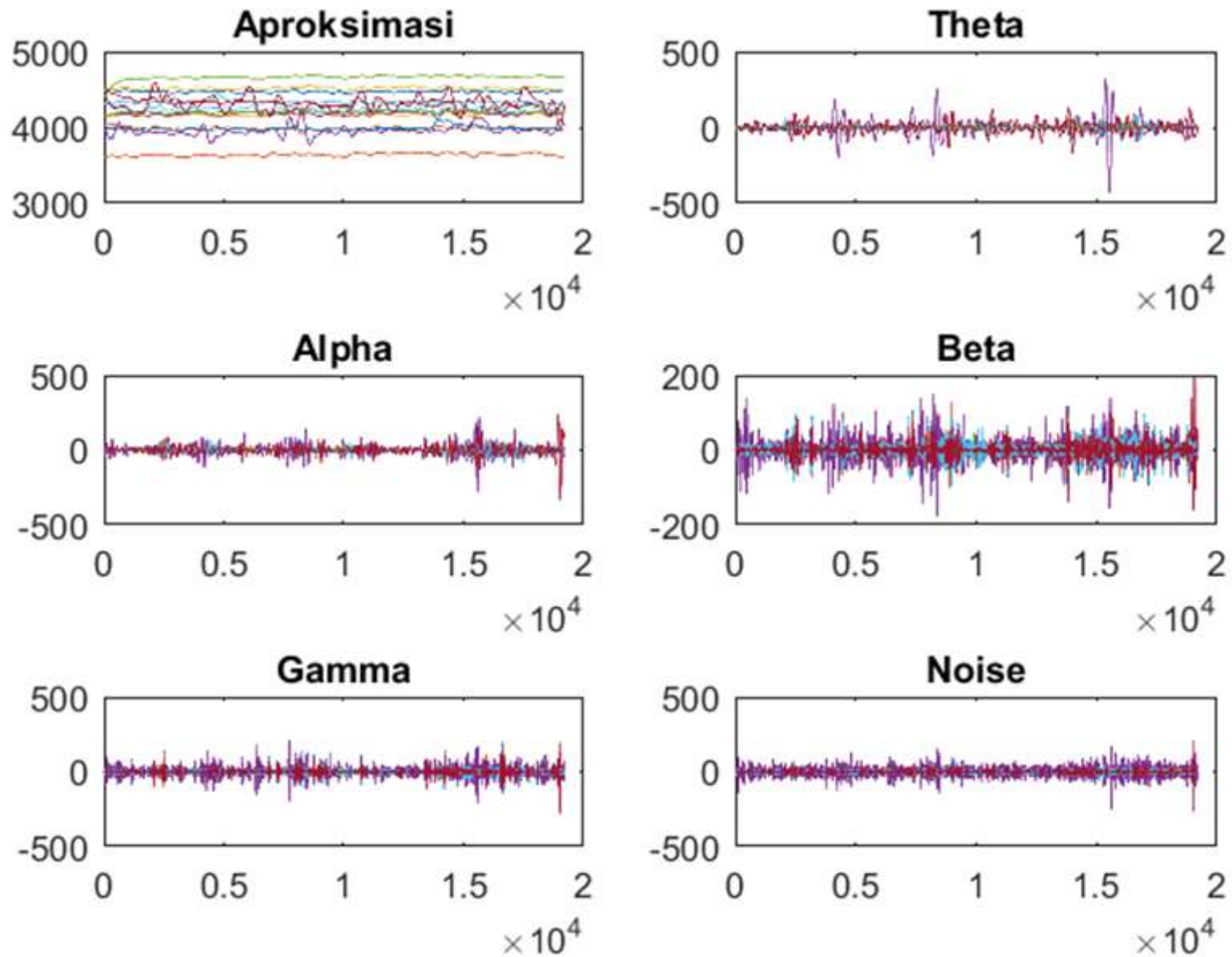
$$W_s f(x) = f(x) * \Psi_s(x) = \frac{1}{s} \int_{-\infty}^{+\infty} f(t) \Psi\left(\frac{x-t}{s}\right) dt \quad (1)$$

The dyadic WT of a digital signal $f(n)$ can be calculated with Mallat Algorithm as follows:

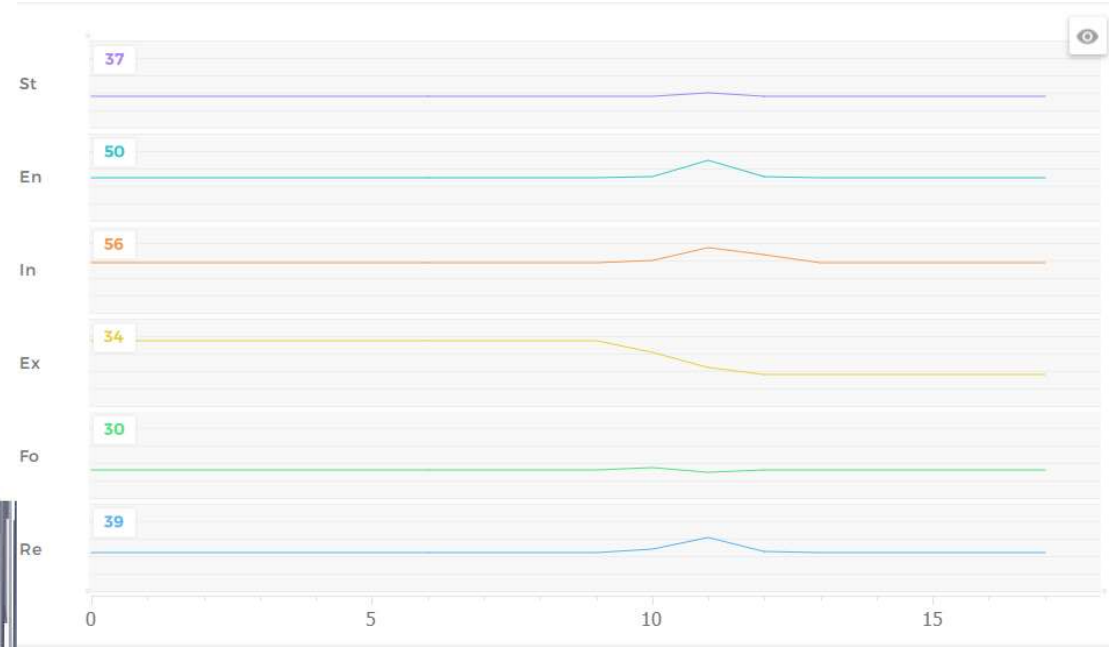
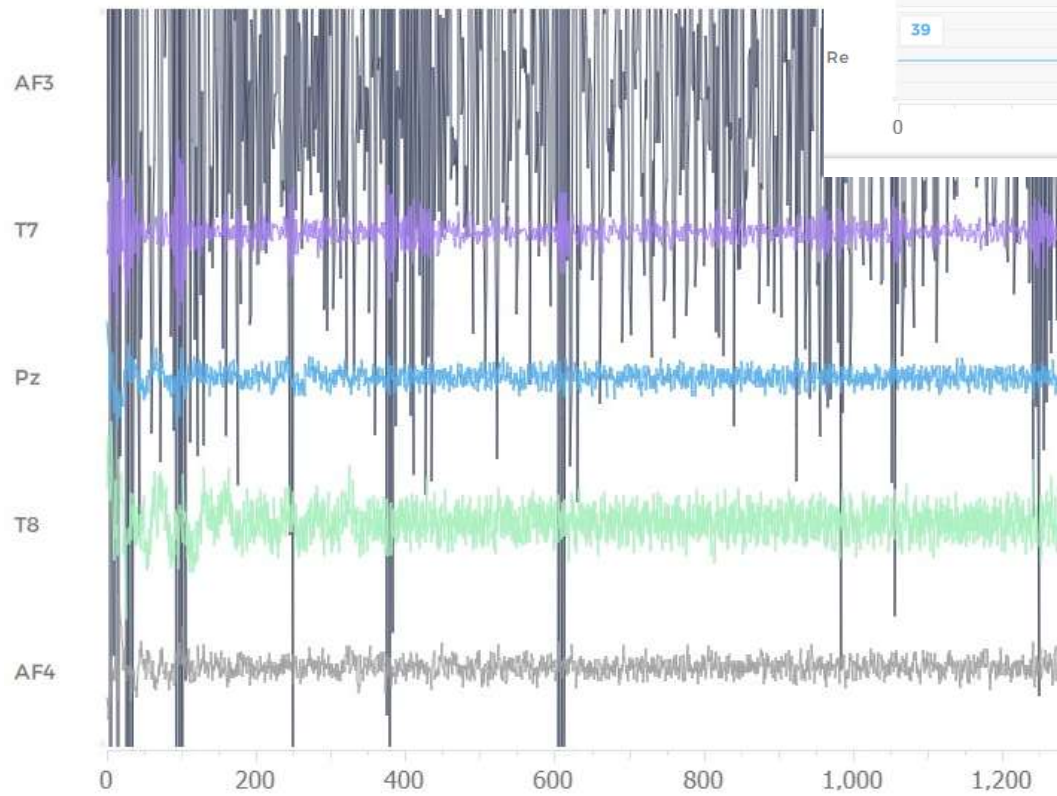
$$S_{2^j} f(n) = \sum_{k \in \mathbb{Z}} h_k S_{2^{j-1}} f(n - 2^{j-1} k) \quad (2)$$

$$W_{2^j} f(n) = \sum_{k \in \mathbb{Z}} g_k S_{2^{j-1}} f(n - 2^{j-1} k) \quad (3)$$

Hasil Ekstraksi gelombang Alva-Beta-Theta-Gamma menggunakan Wavelet



EEG RAW DATA



EMOTIONAL CLASSIFICATION

The Performance of EEG Emotional State Classification

Methods	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	Class	Accuracy (%)
SVM C= 5 $\varepsilon = 10^{-6}$	0.988	0.977	0.557	0.988	0.712	0.043	high	55.77
	0.023	0.012	0.609	0.023	0.045	0.043	low	
	0.558	0.546	0.580	0.558	0.415	0.043	Avg	
RF iteration= 10 bagsize= 100%	0.779	0.745	0.565	0.779	0.655	0.040	high	54.54
	0.255	0.221	0.482	0.255	0.334	0.040	low	
	0.545	0.511	0.528	0.545	0.512	0.040	Avg	
AMGLVQ Iteration = 5 A = 0.6	0.817	0.624	0.640	0.817	0.718	0.215	high	<u>62.98</u>
	0.376	0.183	0.601	0.376	0.462	0.215	low	
	0.630	0.437	0.624	0.630	0.610	0.215	Avg	

Beberapa hal yang masih menjadi tantangan di machine learning

- Data often has poorly understood structure
 - Best modeling approach rarely obvious at start
- Heterogeneous data types
 - e.g. combination of text, images, and numeric data
- Frequent class imbalance
- Noisy or corrupted data
- Missing or incomplete data
- High dimensionality
- Scaling of algorithms to massive data sets
- Streaming (real-time) data

Conclusion and summary

- Beberapa contoh overview dari aplikasi machine learning menunjukkan bahwa pemilihan algoritma sangat berpengaruh pada akurasi sistem.
- Suatu algoritma yang menghasilkan akurasi yang baik pada suatu masalah belum tentu menghasilkan akurasi yang baik juga pada masalah yang lain.
- Percaya atau tidak: in the future, machine learning will play an important role in our daily life.



Terima kasih