Multivariate Classification of Noise Artifacts using Auxiliary Channels - *Artificial Neural Networks*

SangHoon Oh (NIMS)
on behalf of AuxMVC Group in KGWG

Second Korea-Japan Workshop on KAGRA
May 28, 2012

NIMS
National Institute
for Mathematical Sciences
South Korea

Multivariate Classification of Noise Artifacts using Auxiliary Channels - *Artificial Neural Networks*

SangHoon Oh (NIMS)
on behalf of AuxMVC Group in KGWG

Second Korea-Japan Workshop on KAGRA
May 28, 2012

National Institute
for Mathematical Sciences
South Korea

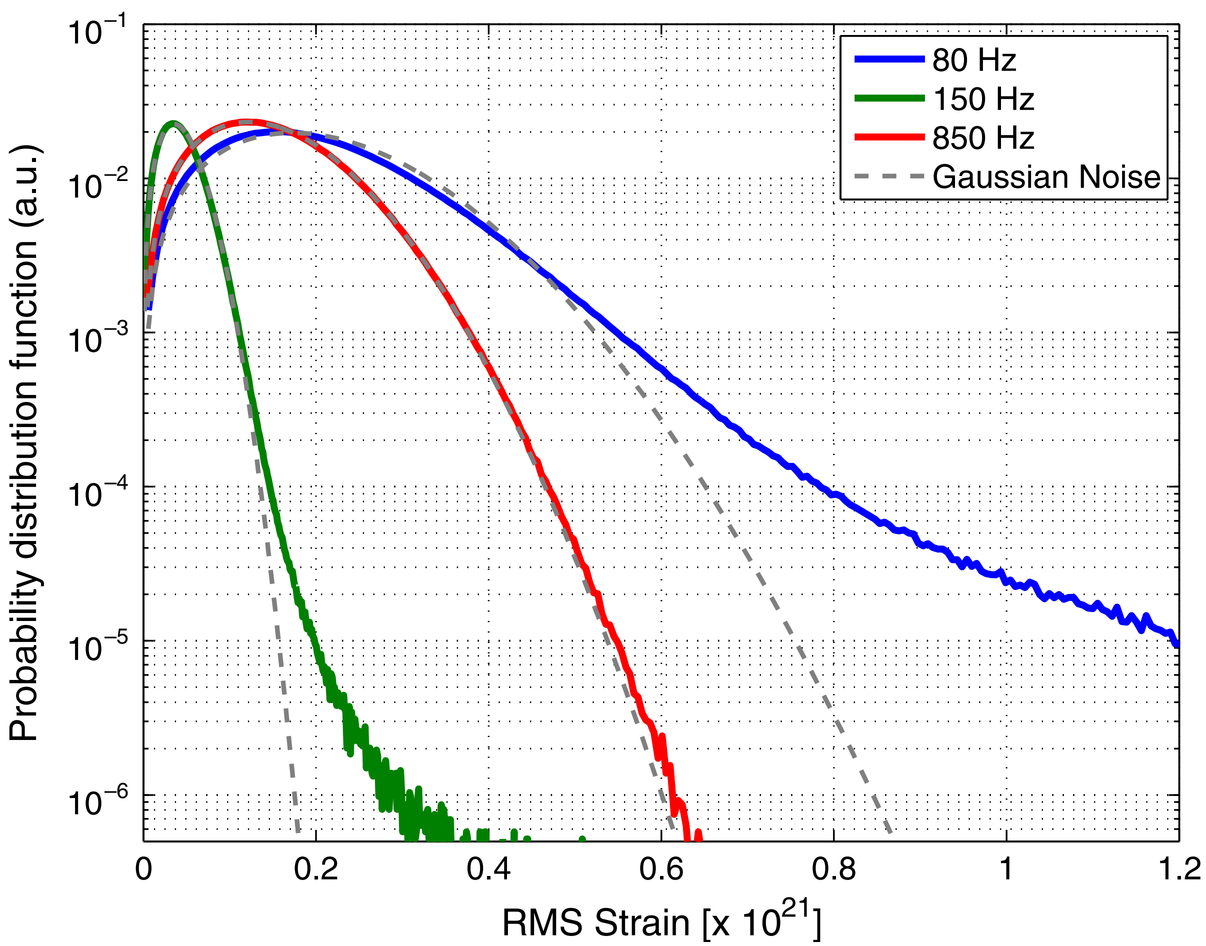
People

- Faculties
  - Sang Hoon Oh (吳 尚勳, NIMS)
  - John J. Oh (吳 廷根, NIMS)
  - Chang-Hwan Lee (Pusan Nat'l Univ.)
- Research Fellow
  - Edwin J. Son (孫 在主,NIMS)
- Grad student
  - Young-Min Kim (金 永敏, NIMS, Pusan Nat'l Univ.)



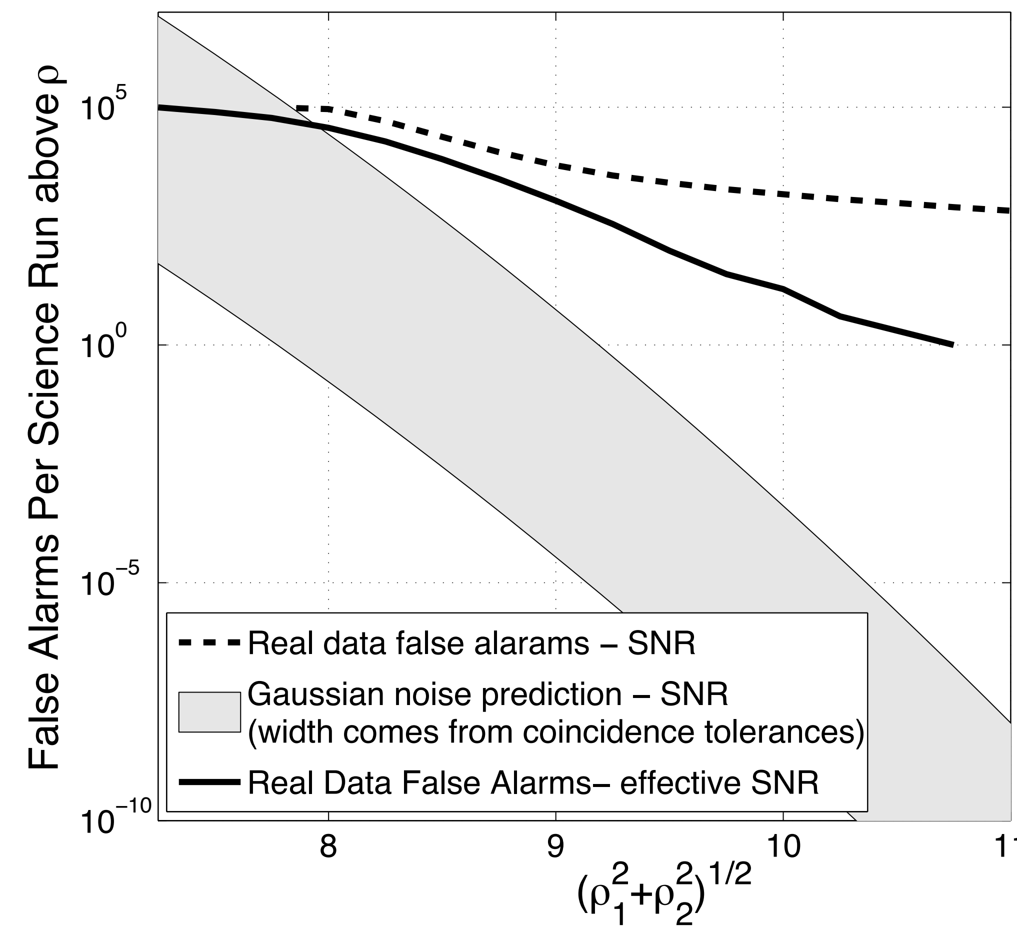

Motivation

Distribution of strain noise amplitude for H1



Abbott et al. (2009)

glitch rate

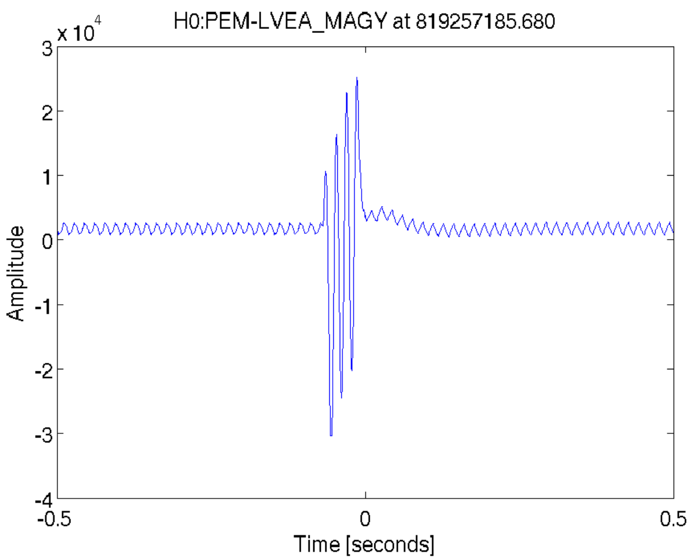


Goals

- Automate process of glitch identification
  - machine learning algorithms
  - $O(10^3)$ of auxiliary channels
- Develop monitoring tools
  - detector characterization
  - feedback to instrumental scientists
  - exploration of important channels in glitch identification

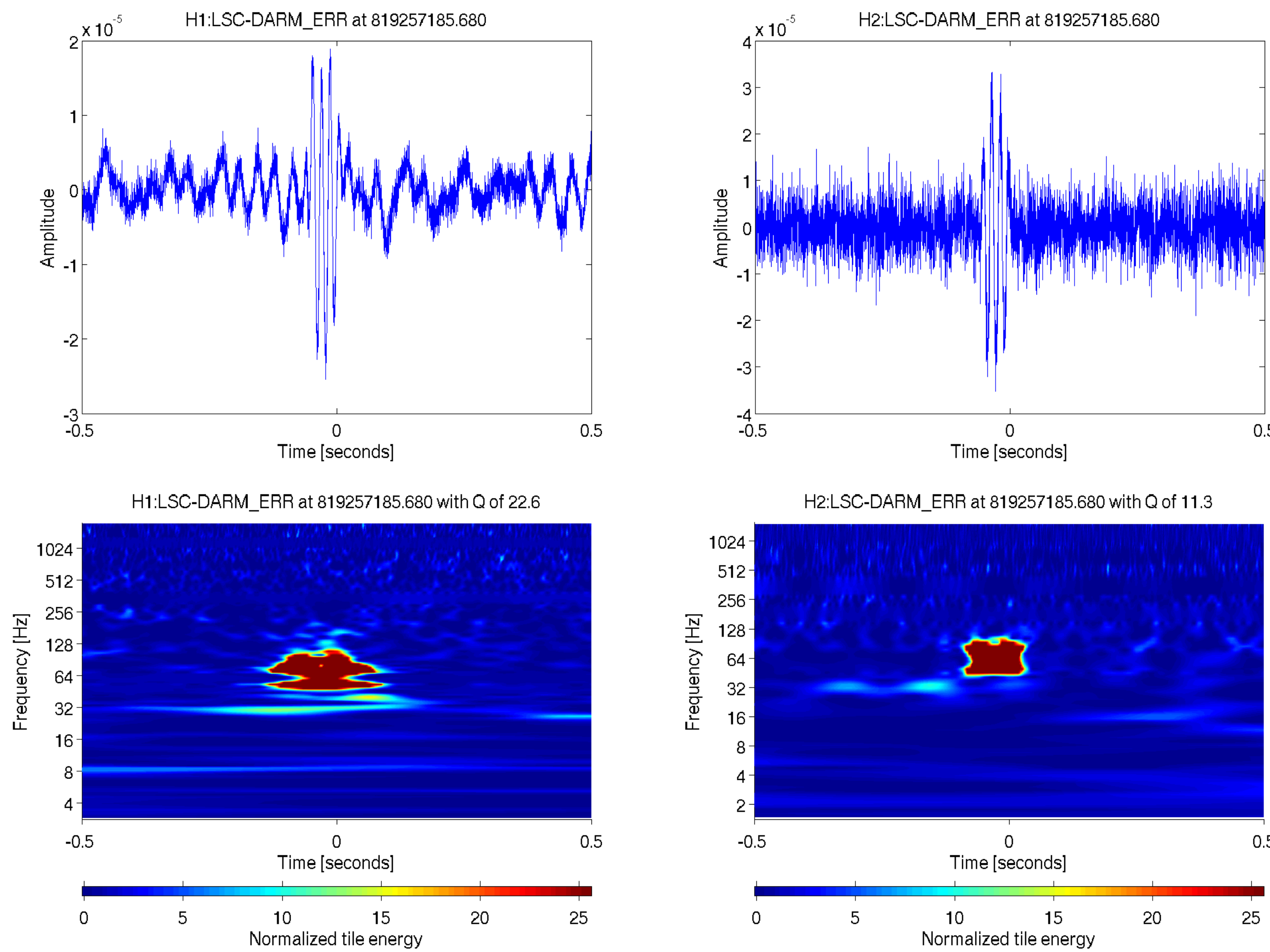
Auxiliary Channel

Power line glitch (noise transient)



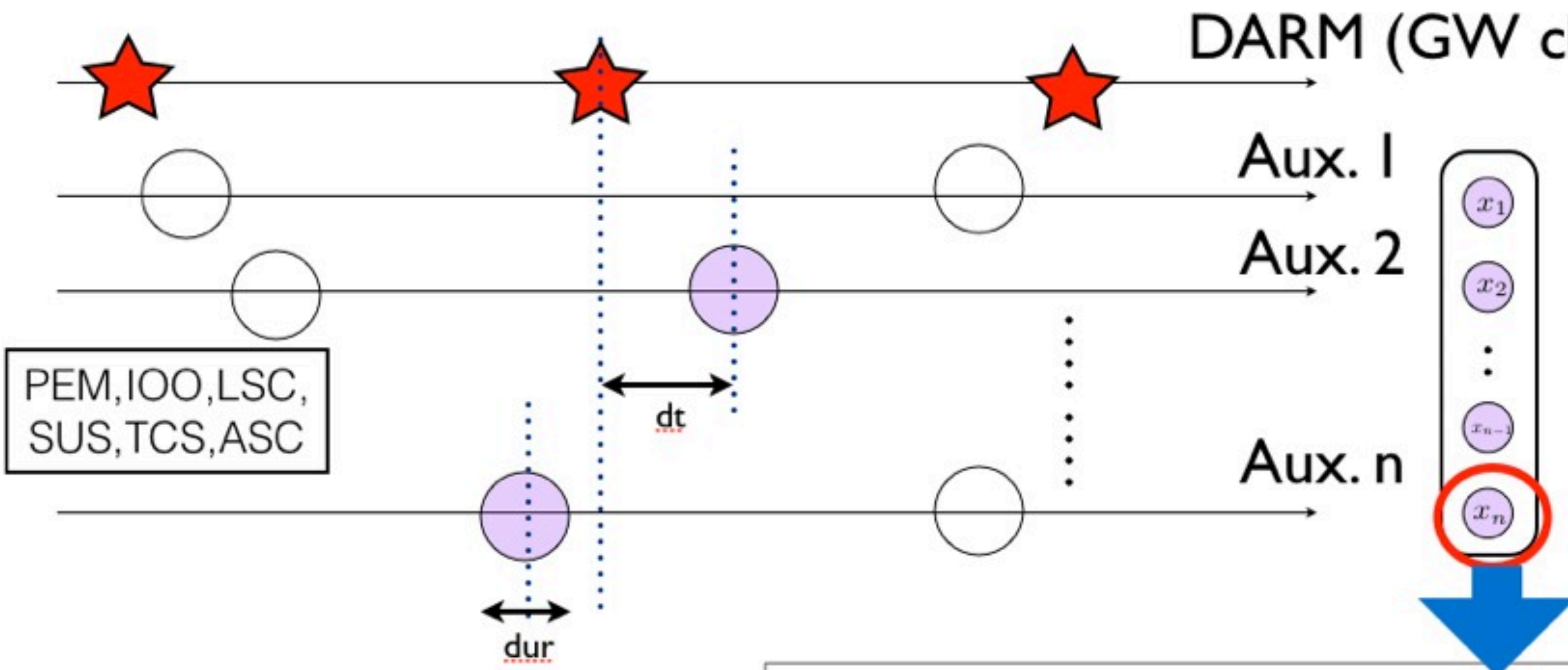
Magnetometer channel

GW channel



- initial LIGO ~1,000 aux channels
- advanced LIGO ~10,000 aux channles

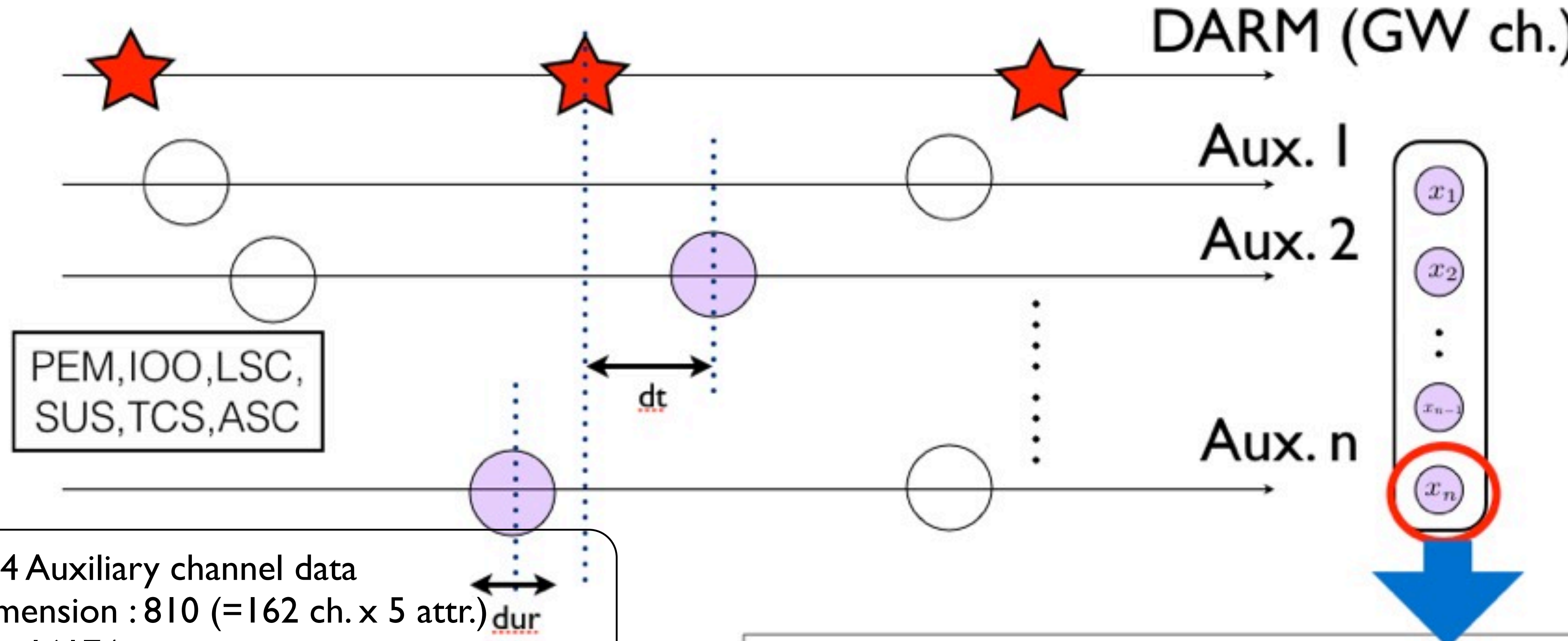
Input Variables



(a) cluster significance of loudest trigger within +/-100ms, or significance of nearest trigger if none exists (significance is related to the SNR)
(b) dt between DARM trigger and AUX trigger in (a)
(c) duration of AUX trigger in (a)
(d) frequency of AUX trigger in (a)
(e) npts (cluster size) of AUX trigger in (a)

signal (=glitch)	clean
loud glitches	random time

Input Variables



LIGO S4 Auxiliary channel data
input dimension : 810 (=162 ch. x 5 attr.)
Glitches : 16174
Clean Samples : 98147

LIGO a week data of S6
(959131741~)
input dimension : 1250
(=250 ch. x 5 attr.)
Glitches : 2826
Clean Sample : 99869

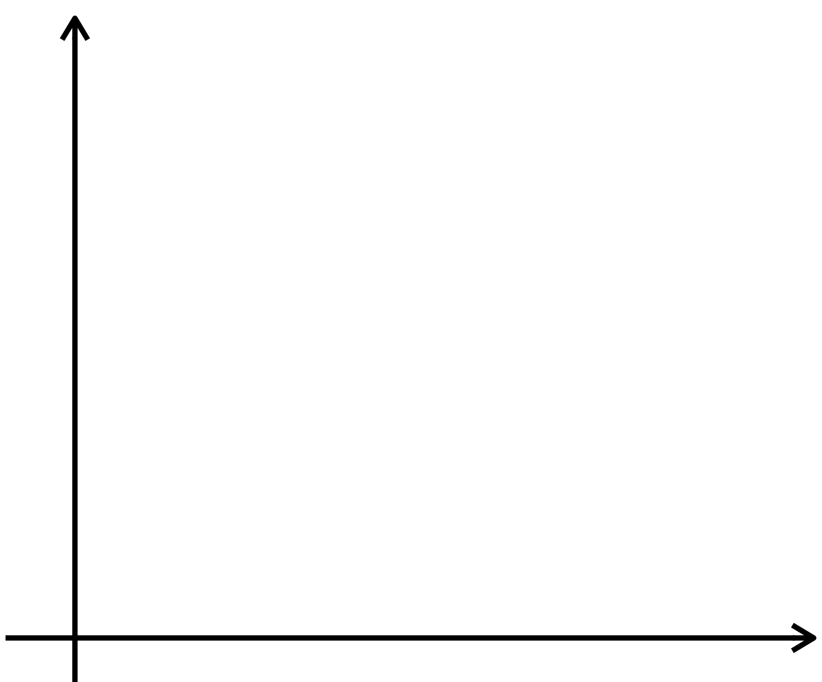
(a) cluster significance of loudest trigger within +/-100ms, or significance of nearest trigger if none exists (significance is related to the SNR)
(b) dt between DARM trigger and AUX trigger in (a)
(c) duration of AUX trigger in (a)
(d) frequency of AUX trigger in (a)
(e) npts (cluster size) of AUX trigger in (a)

Machine Learning Algorithms

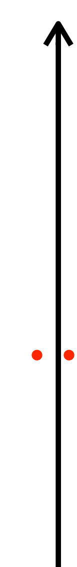
- Algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or database.
- Automatic learning to recognize complex patterns and make intelligent decision based on data

(from wikipedia)

Stanford online ML Class



classification



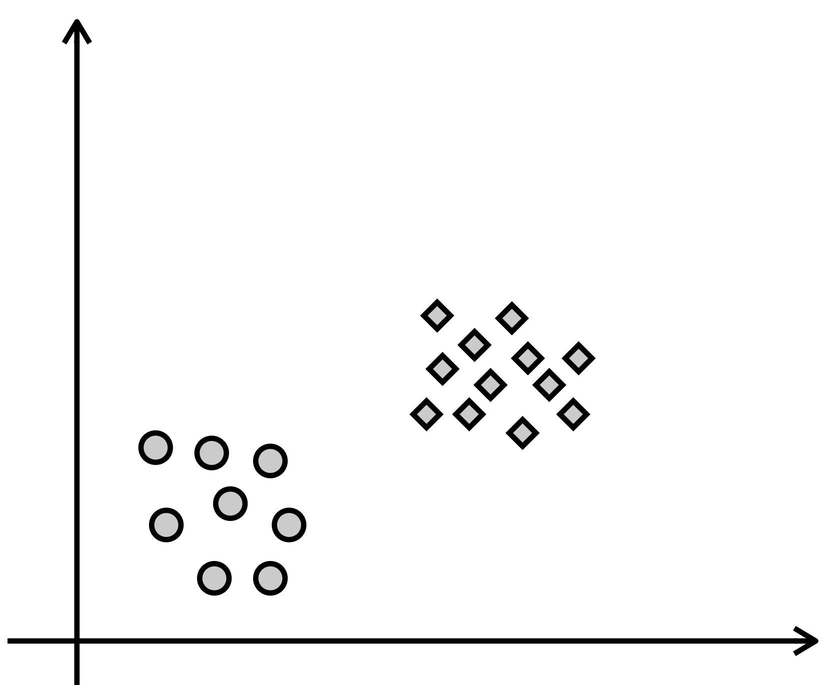
Machine Learning Algorithms

- Algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or database.

- Automatic learning to recognize complex patterns and make intelligent decision based on data

(from wikipedia)

Stanford online ML Class



classification

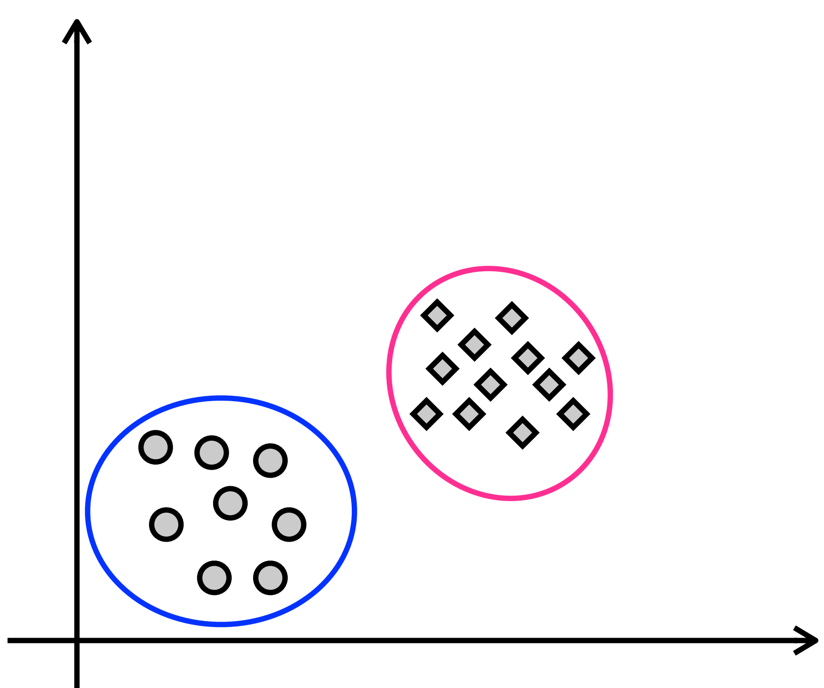
Machine Learning Algorithms

- Algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or database.

- Automatic learning to recognize complex patterns and make intelligent decision based on data

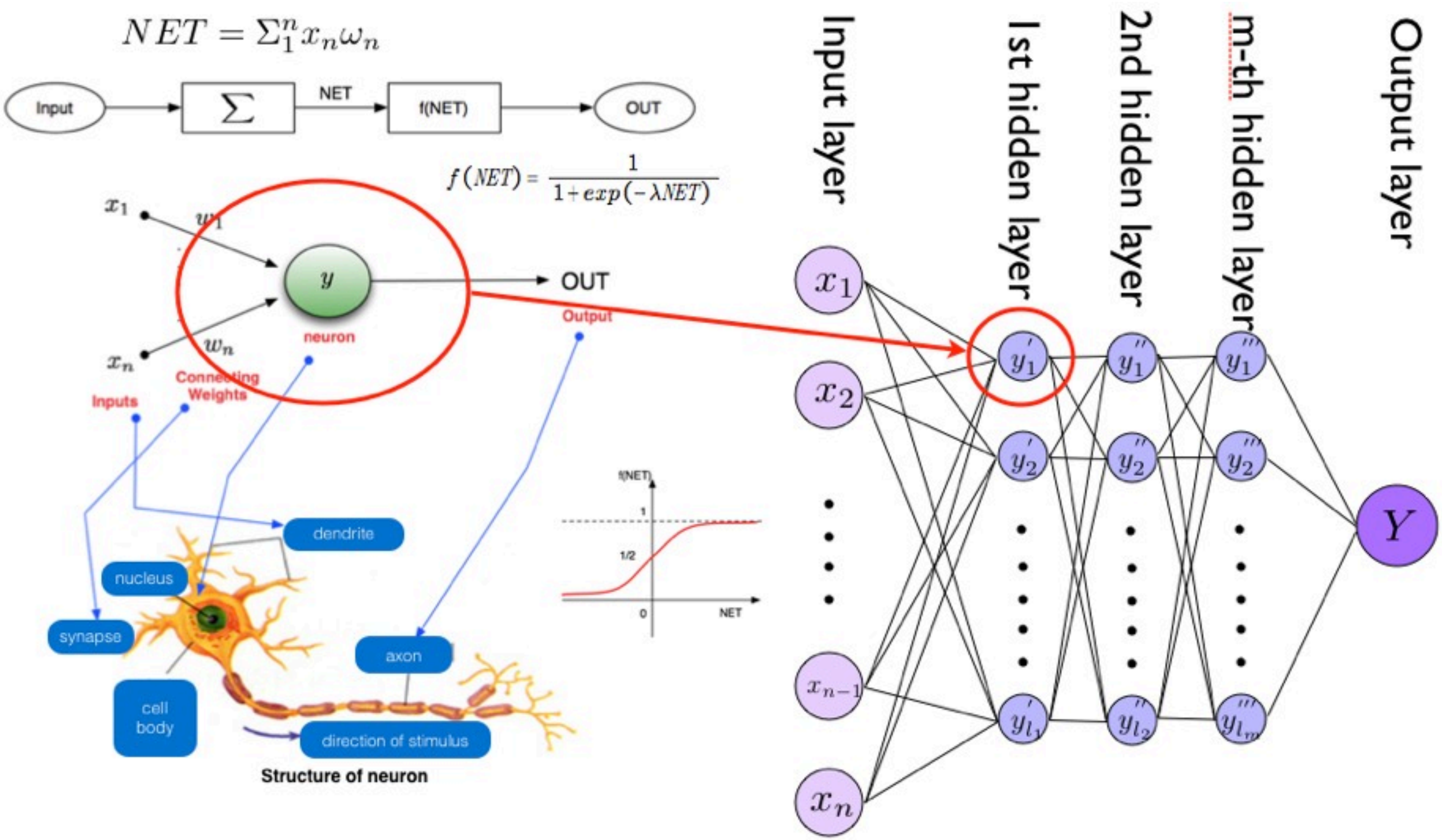
(from wikipedia)

Stanford online ML Class



classification

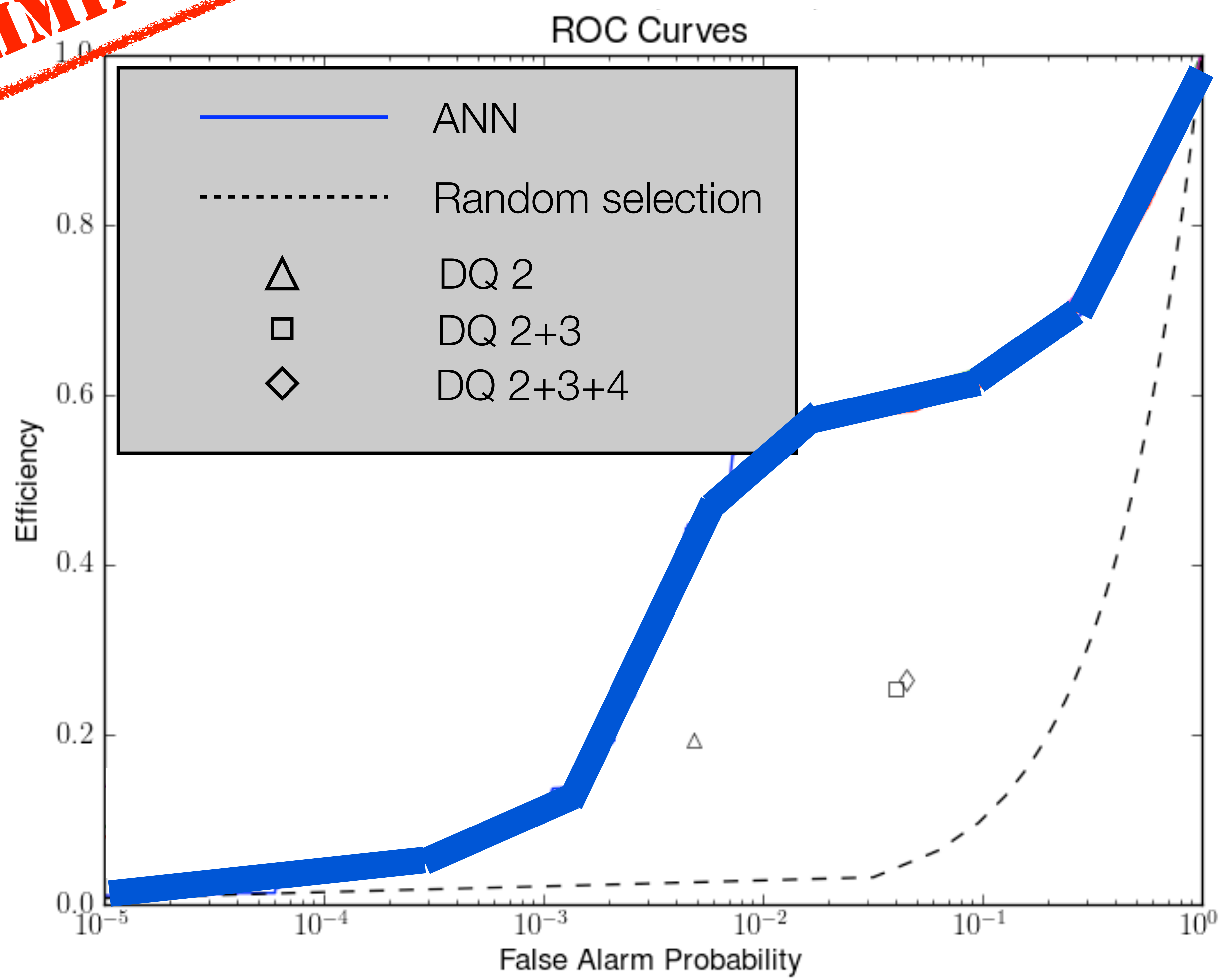
Artificial Neural Networks



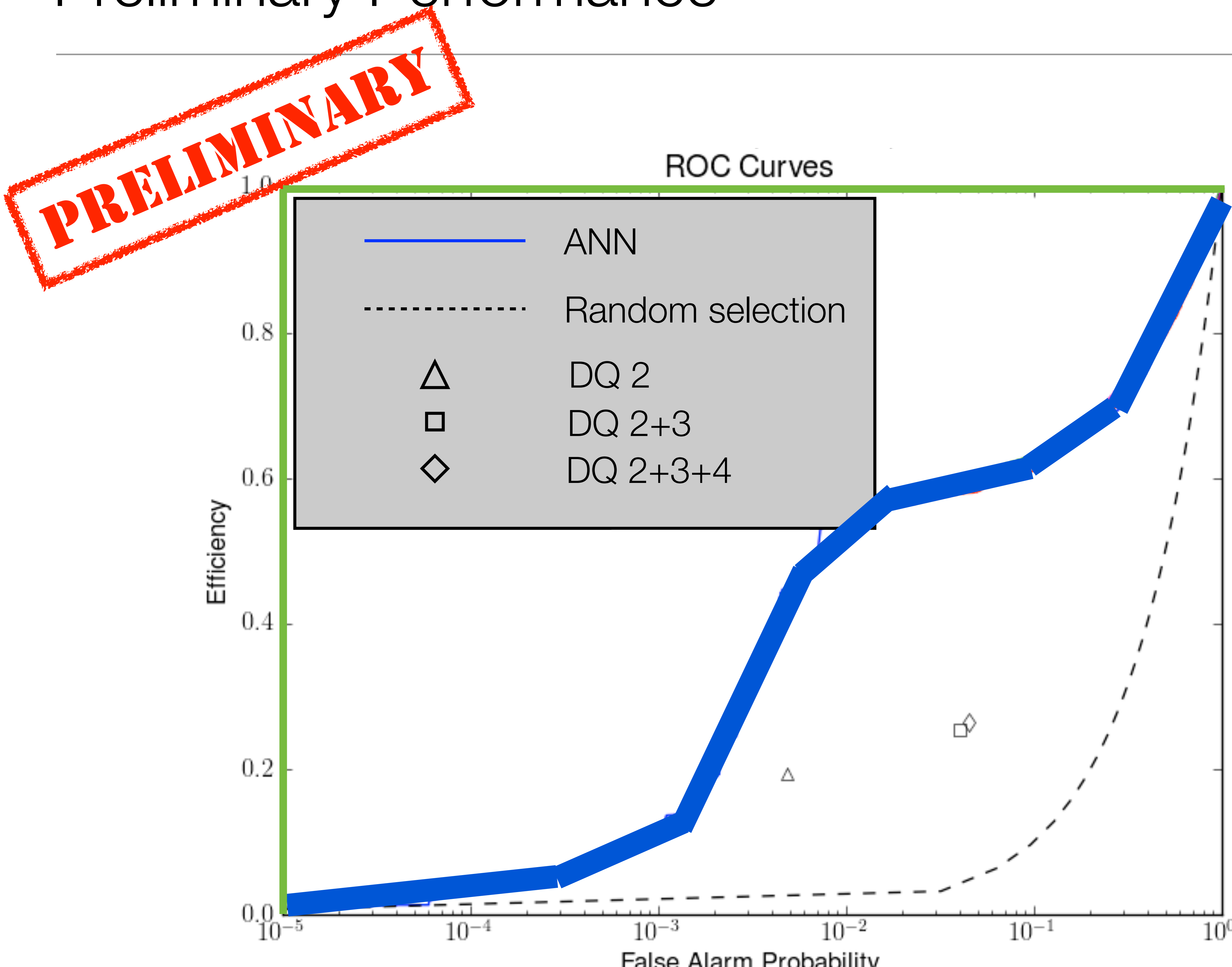
KAGRA

Preliminary Performance

PRELIMINARY



Preliminary Performance

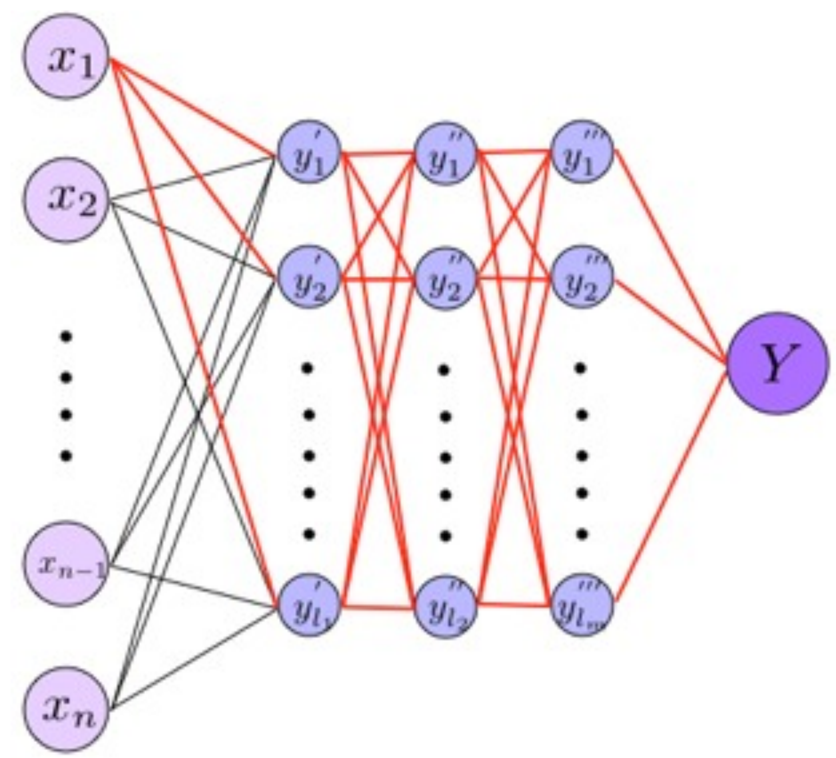


Important Channels

- Significant Factor (SF)
  - Indicate an quantitative importance of an input variable(***x**_i*) to give an output(***Y***)
  - In ANN, total summation of all connection weights(red lines) connected to a given input variable(***x**_i*) is considered as SF.

$$S(i) = \Sigma_M |(w(i,l_1)w(l_1,l_2)....w(l_{M-1},l_M)w(l_M,o)|$$

$$\Sigma_M = \Sigma_{l_1=1}^{L_1}\Sigma_{l_2=1}^{L_2}.....\Sigma_{l_{M-1}=1}^{L_{M-1}}$$



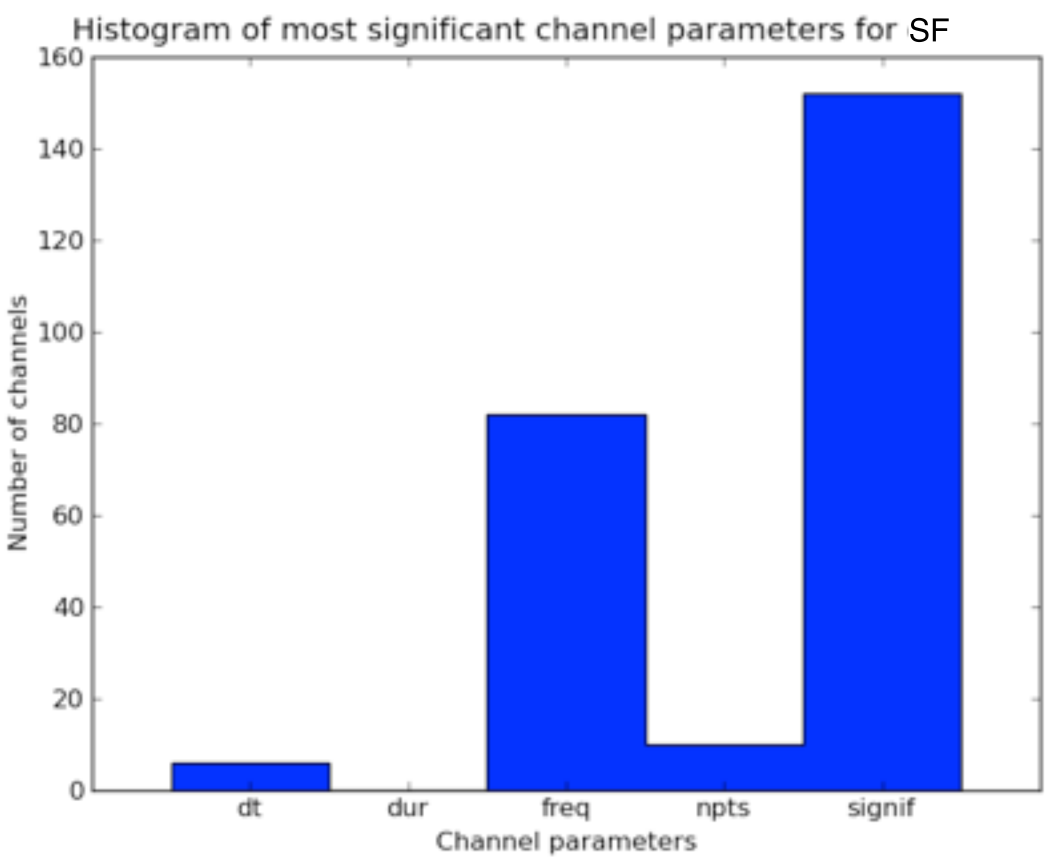
- The larger SF indicates the stronger association with noise artifacts in GW channel.

Important Channels

PRELIMINARY

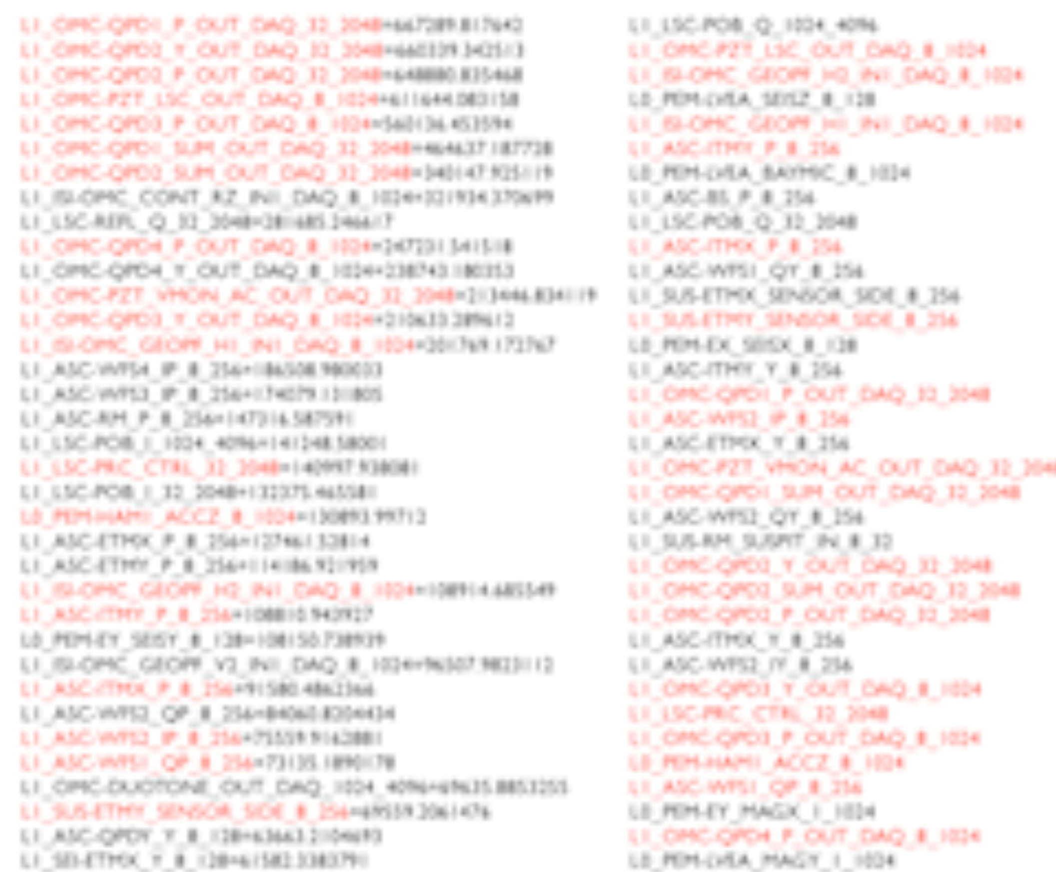
Test week

- 19 of top 35 channels in ANN are matched with the result of HVeto.



ANN

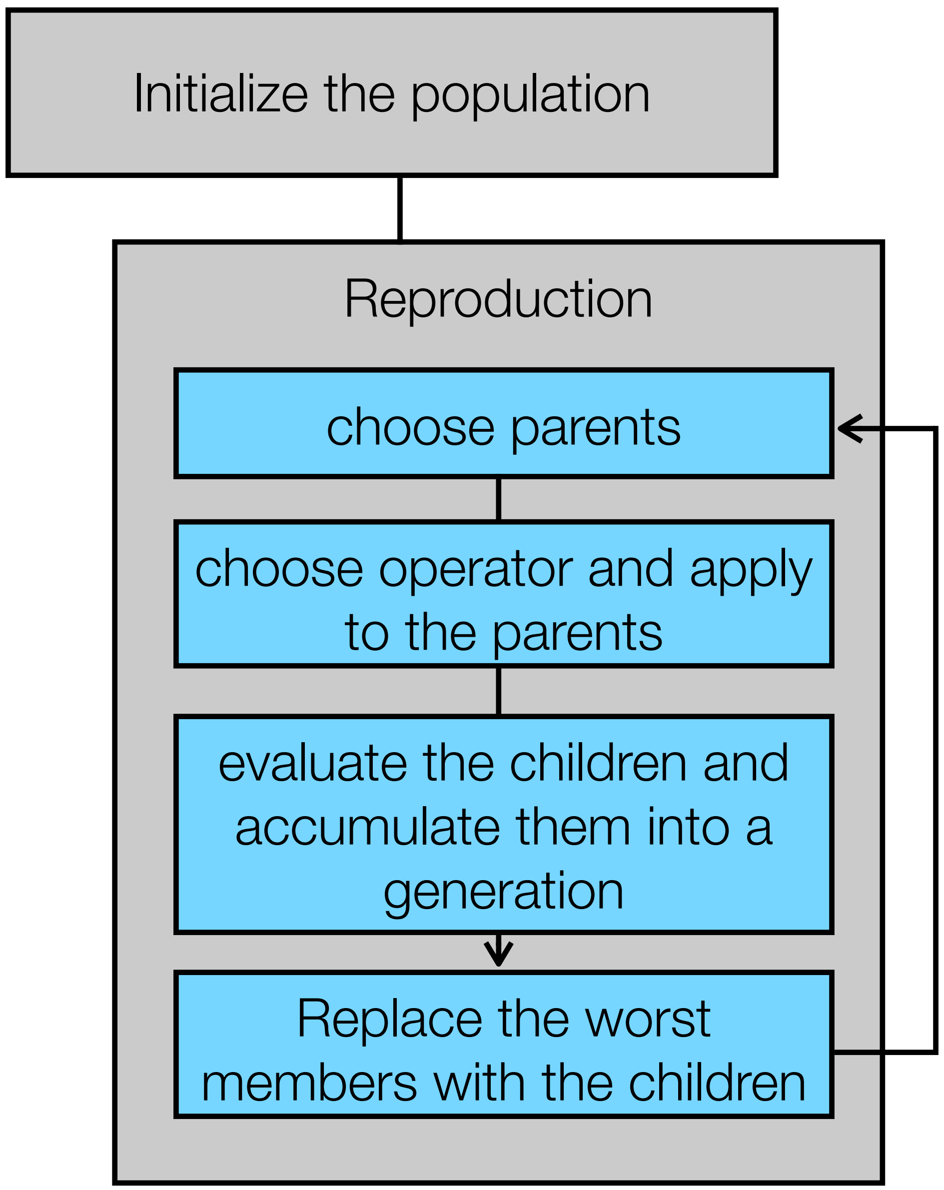
HVeto

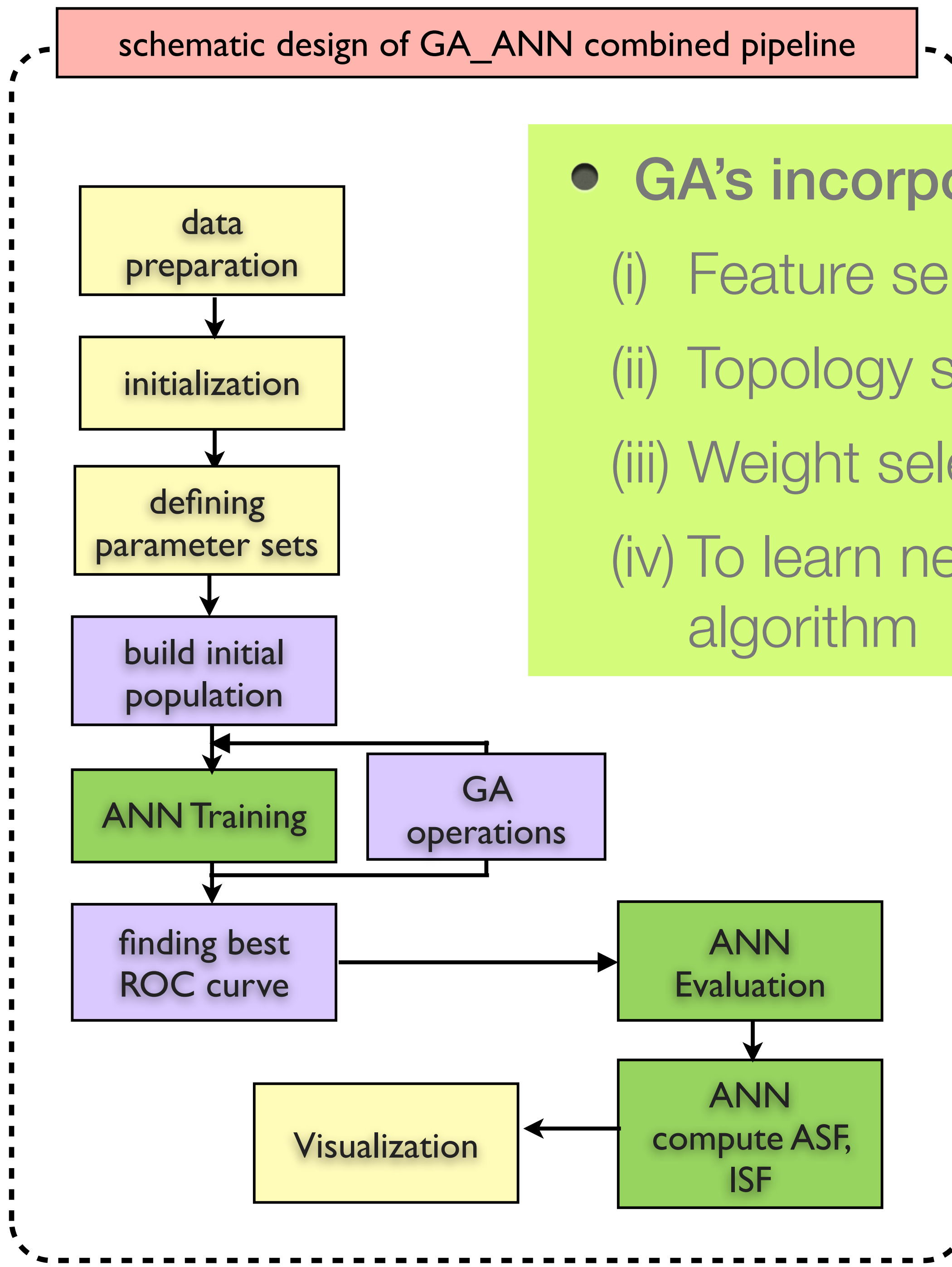


Genetic Algorithm

- Optimization / Machine learning (loosely) based on biological evolution; ***natural selection of genes***

Search global optimum





- **GA's incorporation with ANN**
  - (i) Feature selection
  - (ii) Topology selection
  - (iii) Weight selection
  - (iv) To learn neural network learning algorithm

Summary

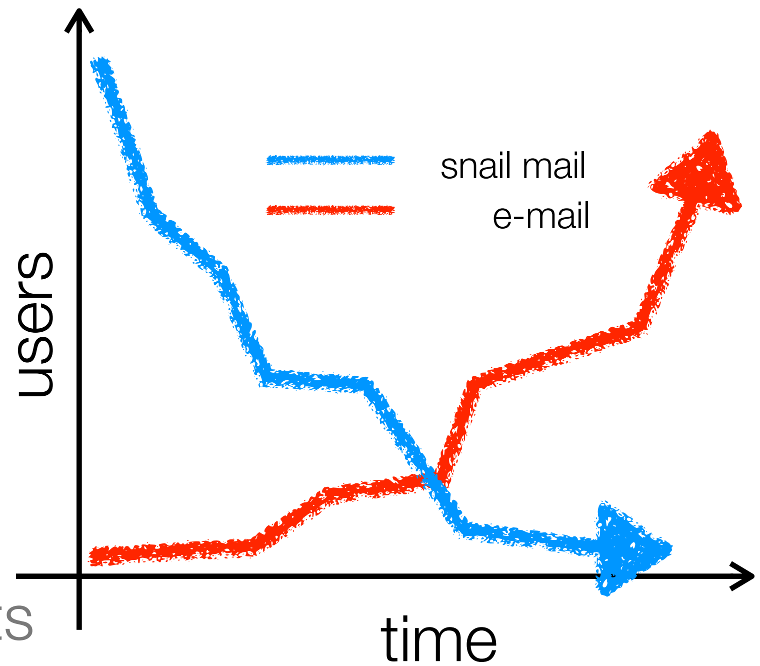
- Artificial Neural Networks (ANNs) has been applied to auxiliary channels to identify noise artifacts in the GW channel.
- Handling 810 input variables per trigger, feeding into ANN's input layer.
- Preliminary results show better performance than DQ category.
- Genetic algorithm can optimize ANN for choosing the initial guess of connection weights, input features, and values of topological parameters.
- Worth further investigation

Thank you for your attention!

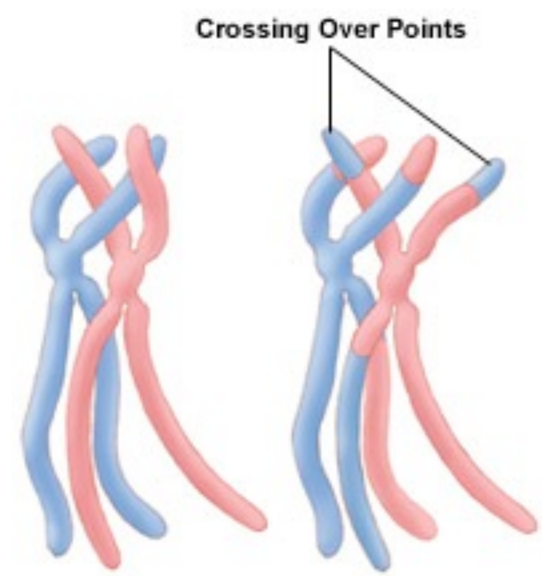
ありがとうございます。

Genetic Algorithm

- Optimization / Machine learning (loosely) based on biological evolution; ***natural selection of genes***



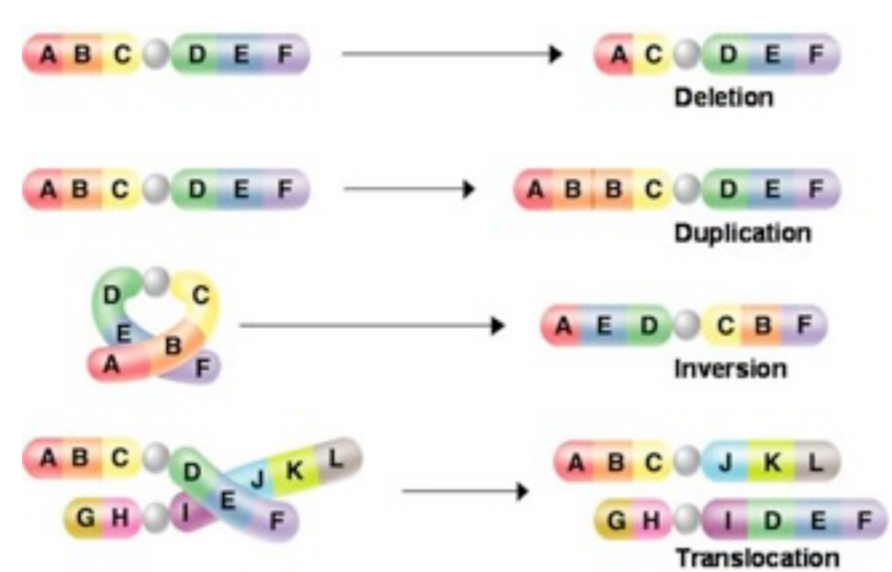
crossover



- Five components
  - (i) encoding solutions into chromosomes (**string**)
  - (ii) an evaluation function
  - (iii) initialization of population of chromosomes
  - (iv) operators for reproduction: mutation and crossover
  - (v) parameter settings for the algorithm, the operators, i.e., *N*(population), *N*(generations), *prob*(operators)

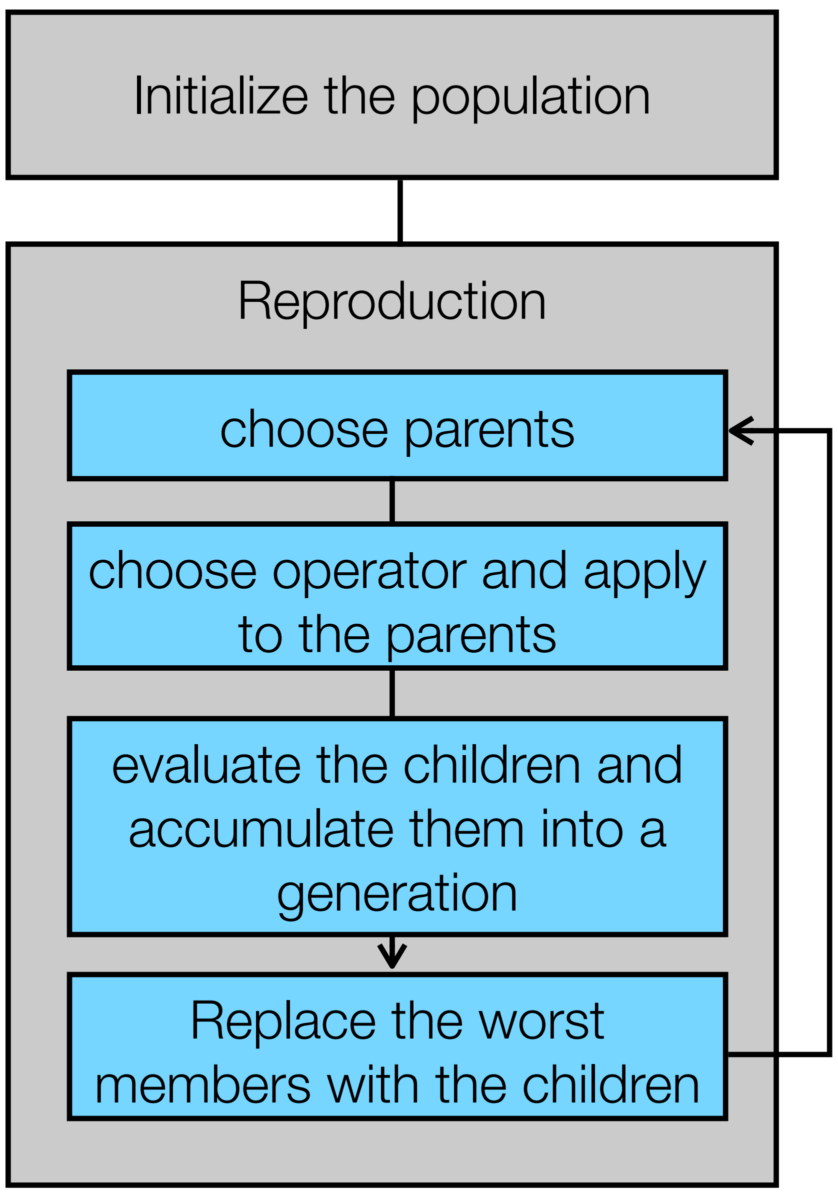
Text

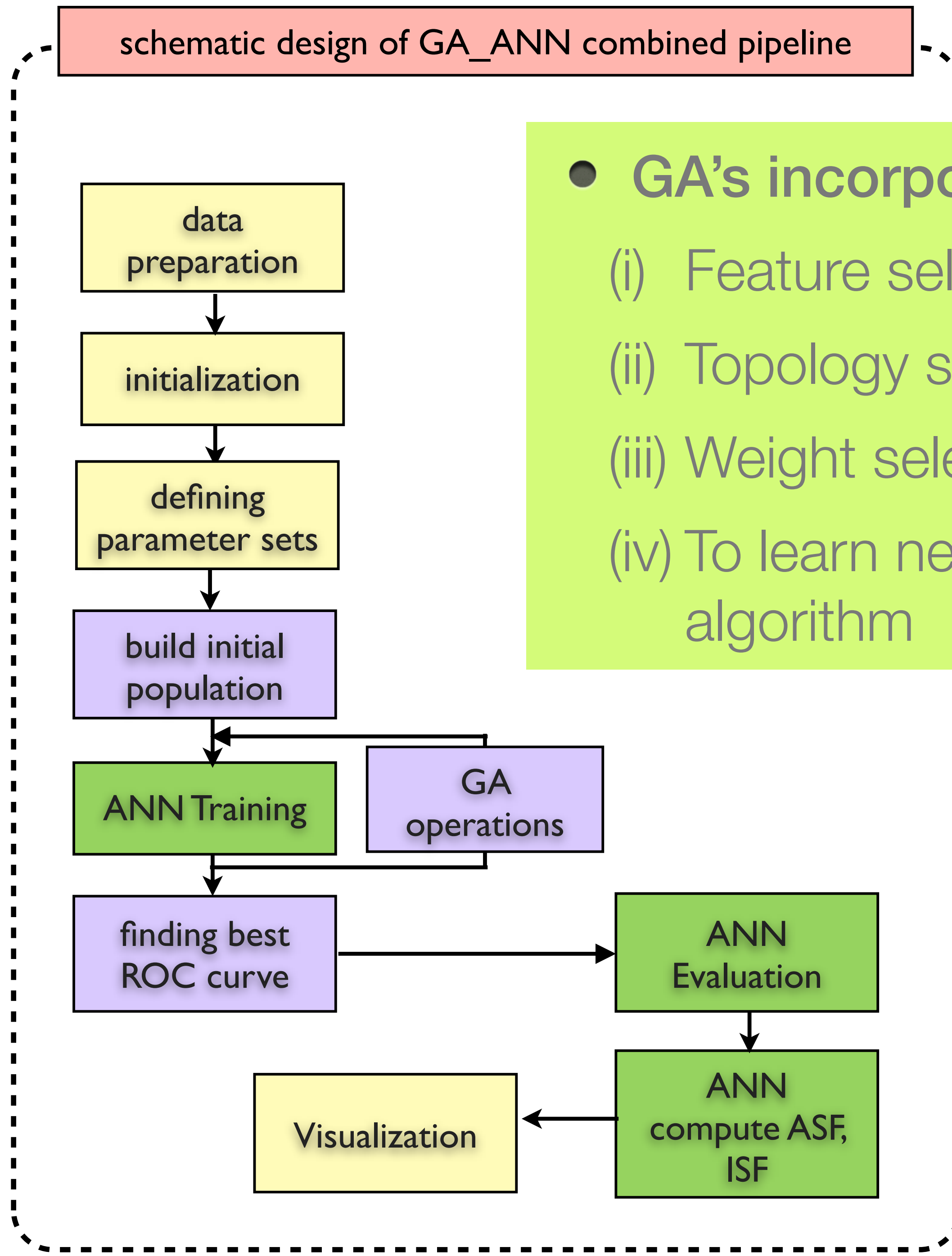
mutation



Genetic Algorithm

- Operation of GA

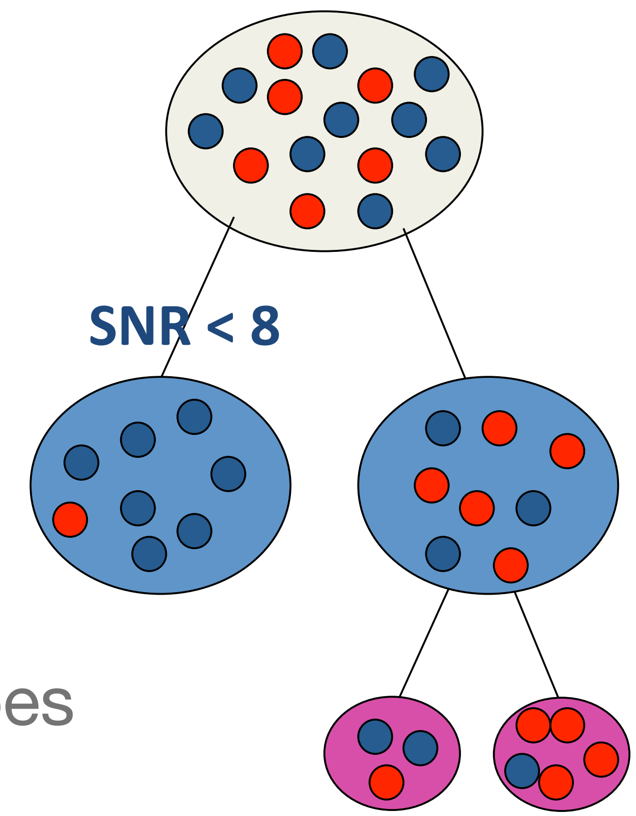


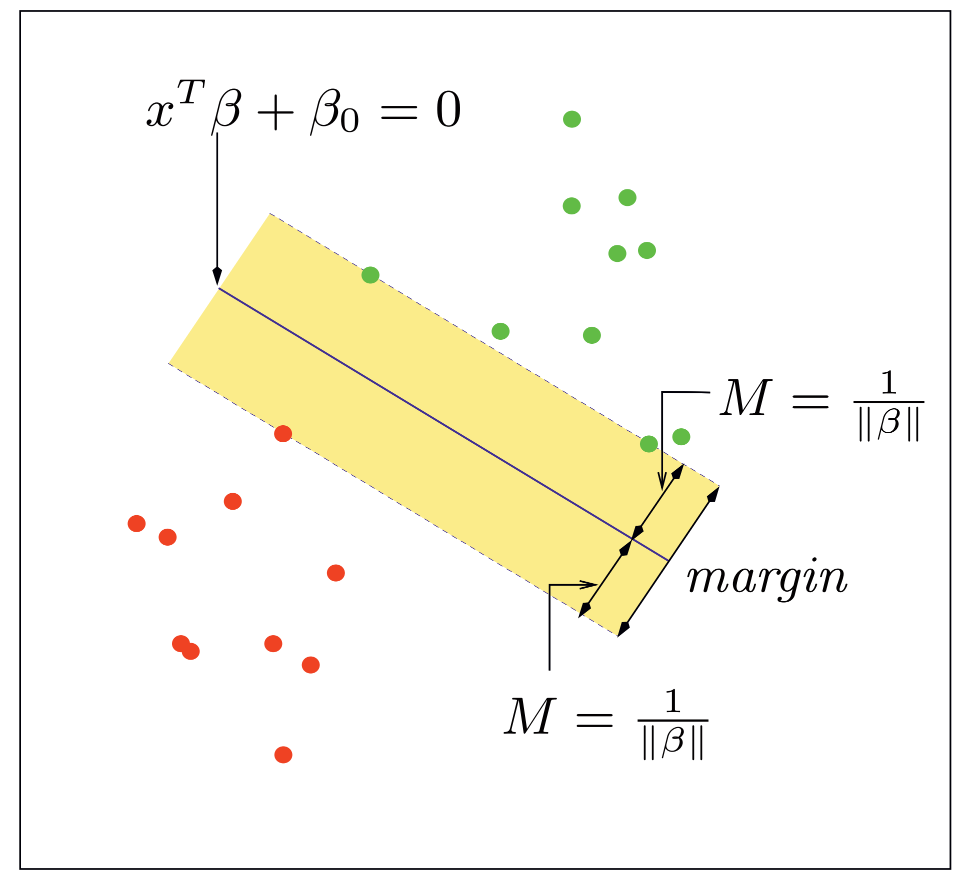


- **GA's incorporation with ANN**
  - (i) Feature selection
  - (ii) Topology selection
  - (iii) Weight selection
  - (iv) To learn neural network learning algorithm

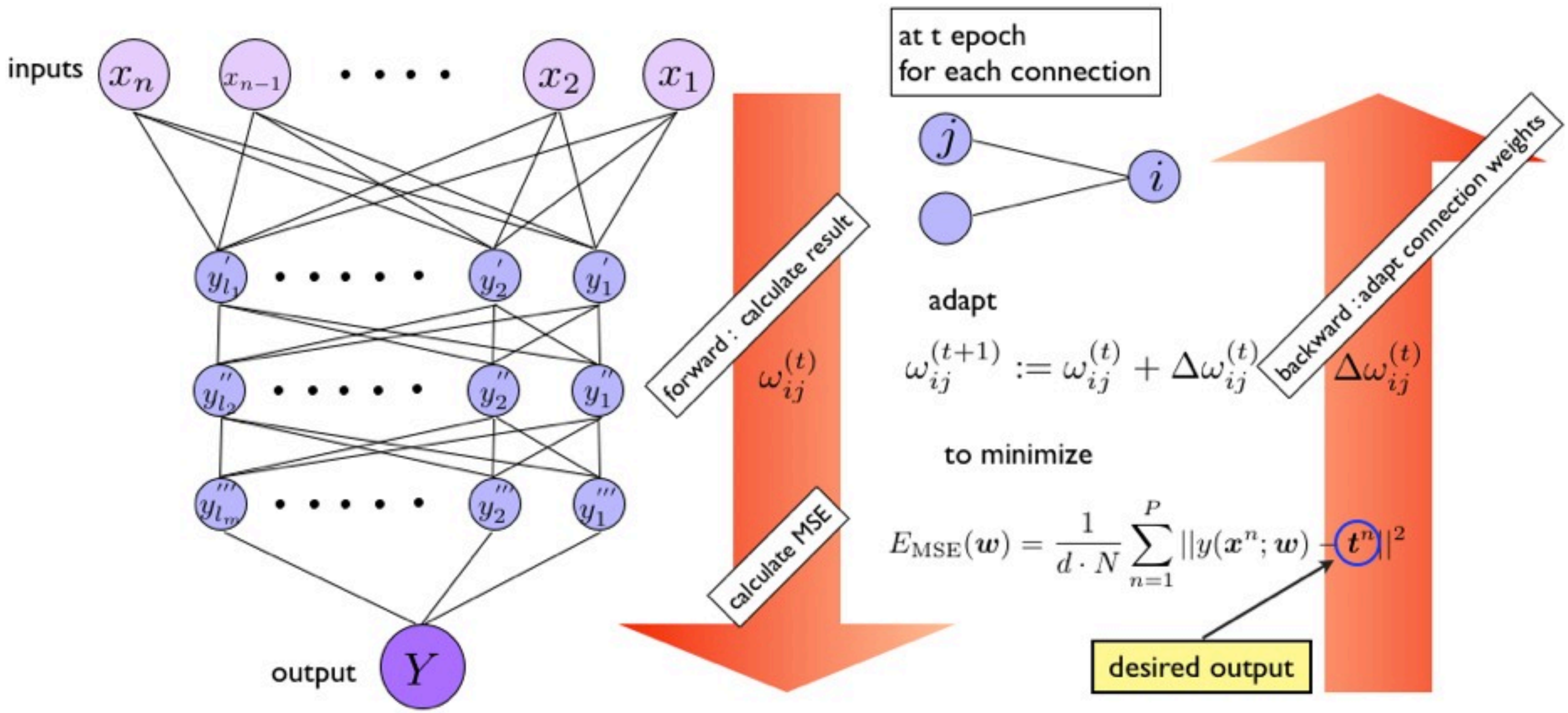
Machine Learning Algorithms in practice

- MVSC
- SVN
- ANN
- Hierarchical vetoes





Artificial Neural Networks



Artificial Neural Networks

iRPROP(Igel &Hüsken,2009)
implemented in FANN lib.

$$\omega_{ij}^{(t+1)} := \omega_{ij}^{(t)} + \Delta\omega_{ij}^{(t)}$$

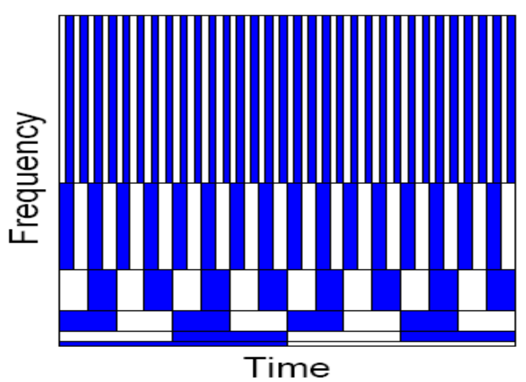
$$\Delta w_{ij}^{(t)} := -\operatorname{sign}\left(\frac{\partial E}{\partial w_{ij}}^{(t)}\right) \cdot \Delta_{ij}^{(t)}$$

$$\Delta_{ij}^{(t)} := \begin{cases} \min\left(\eta^{+} \cdot \Delta_{ij}^{(t-1)}, \Delta_{\max}\right) & , \text{ if } \frac{\partial E}{\partial w_{ij}}^{(t-1)} \cdot \frac{\partial E}{\partial w_{ij}}^{(t)} > 0 \\ \max\left(\eta^{-} \cdot \Delta_{ij}^{(t-1)}, \Delta_{\min}\right) & , \text{ if } \frac{\partial E}{\partial w_{ij}}^{(t-1)} \cdot \frac{\partial E}{\partial w_{ij}}^{(t)} < 0 \\ \Delta_{ij}^{(t-1)} & , \text{ otherwise ,} \end{cases}$$

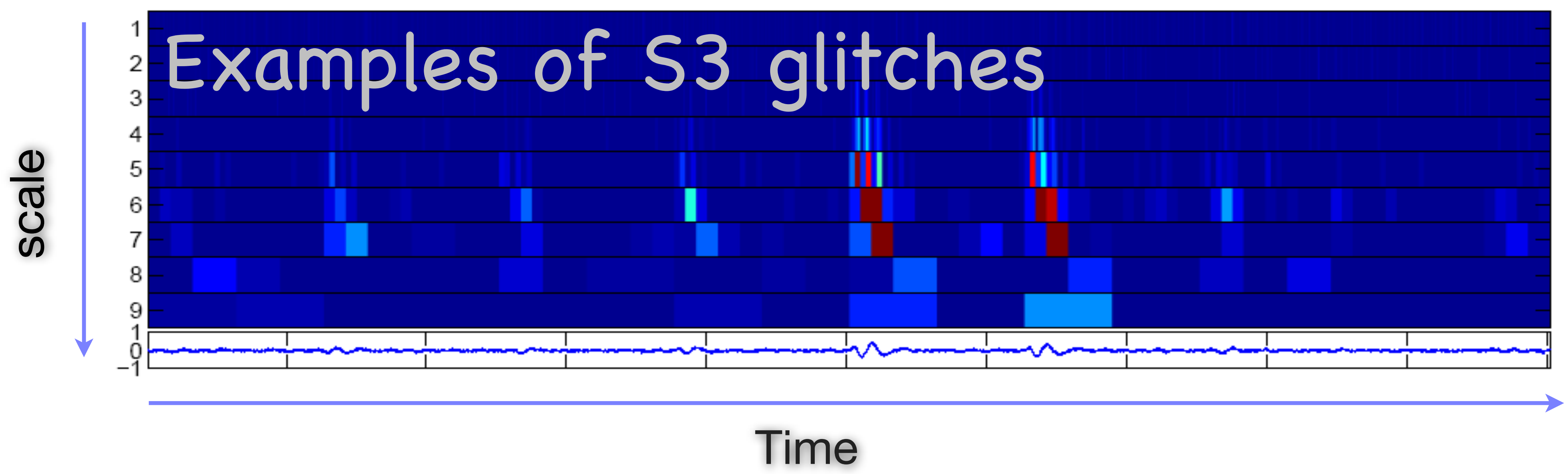
$$\frac{\partial E^{(t)}}{\partial \Delta_{ij}^{(t-1)}} = \frac{\partial E^{(t)}}{\partial w_{ij}^{(t)}} \frac{\partial w_{ij}^{(t)}}{\partial \Delta_{ij}^{(t-1)}} = -\frac{\partial E}{\partial w_{ij}}^{(t)} \operatorname{sign}\left(\frac{\partial E}{\partial w_{ij}}^{(t-1)}\right)$$

Input Variables

KleineWelle Wavelet Transformation



Discrete Dyadic Wavelet transform to decompose the time series into a logarithmically-spaced time-frequency plane



KAGRA

Input Variables

(a) Significance of loudest trigger within +/- 100ms, or significance of nearest trigger if none exists (sig is related to the SNR)

(b) Δt to triggers in (a)

(c) duration of trigger in (a)

(d) freq of AUX trigger in (a)

(e) npts in AUX trigger in (a)

signal (=glitch)	clean
loud glitches	random time