適応システム理論
ガイダンス

Introduction to Neural Networks

Kenji Nakayama
Kanazawa University, JAPAN

Multi-Layer Neural Networks

PPTファイルの入手方法

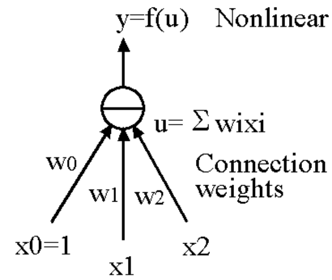
下記URLからダウンロード

http://leo.ec.t.kanazawa-u.ac.jp/~nakayama/edu/neural.htm

質問はメールでお願いします

nakayama@t.kanazawa-u.ac.jp

Artificial Neuron Model



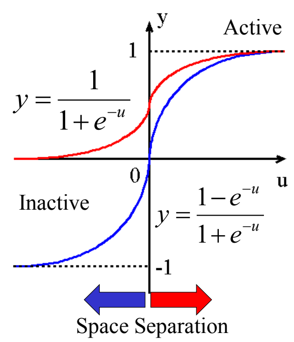
Neural Networks

Network Structures
- Multi-layer Network
- Recurrent Network

Learning Algorithms
- Supervised Learning
- Unsupervised Learning

Functions
- Pattern Mapping and Classification
- Estimation and Prediction
- Associative Memory
- Optimization and Minimization

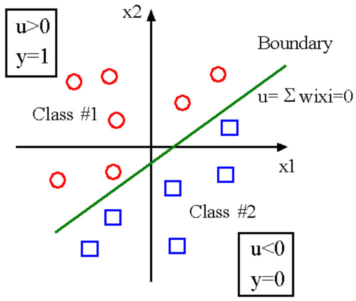
Activation (Nonlinear) Function of Neuron



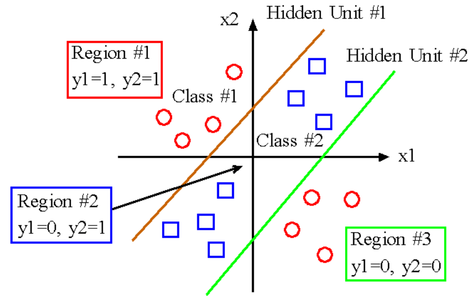
$$y = \frac{1}{1+e^{-u}}$$

$$y = \frac{1-e^{-u}}{1+e^{-u}}$$

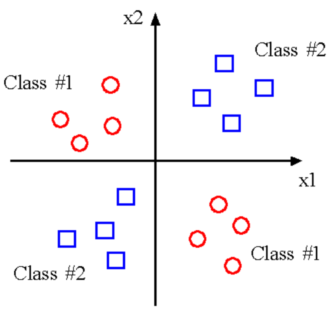
Pattern Classification by Single Neuron



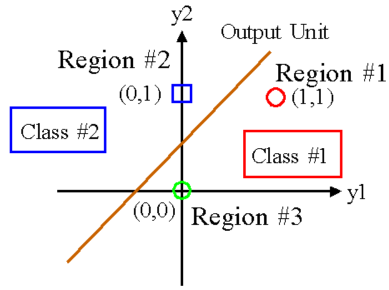
Pattern Classification by Two-Layer NN
- Region Separation by Hidden Units-



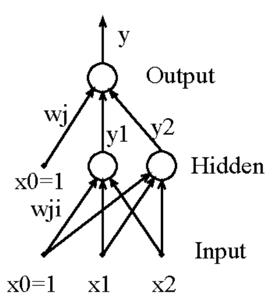
Linearly Inseparable Problem



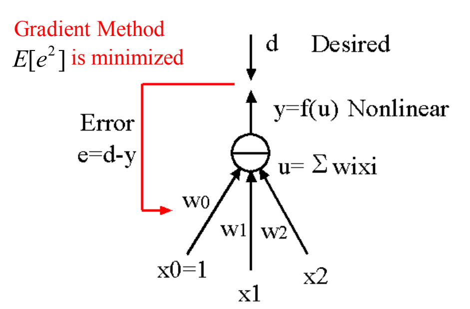
Pattern Classification by Two-Layer NN
- Class Separation by Output Unit -

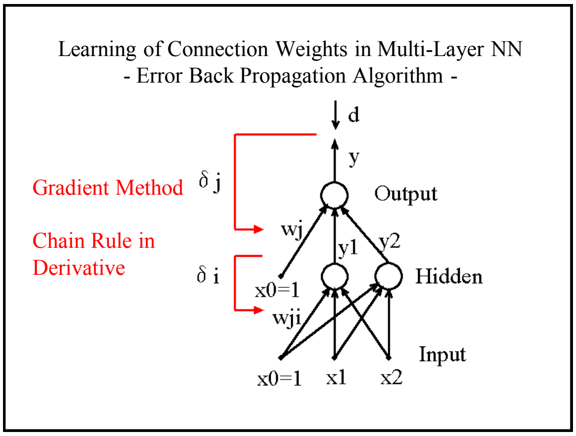


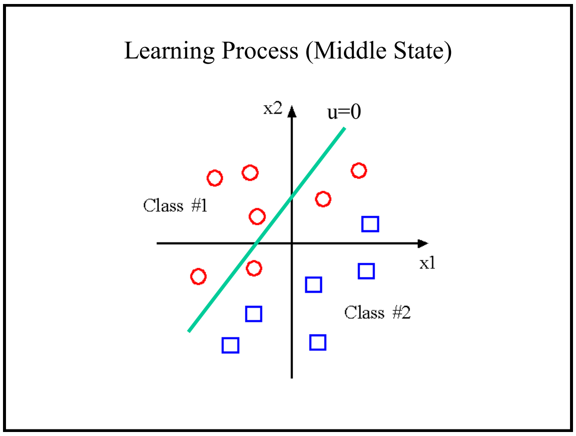
Two Layer Neural Network

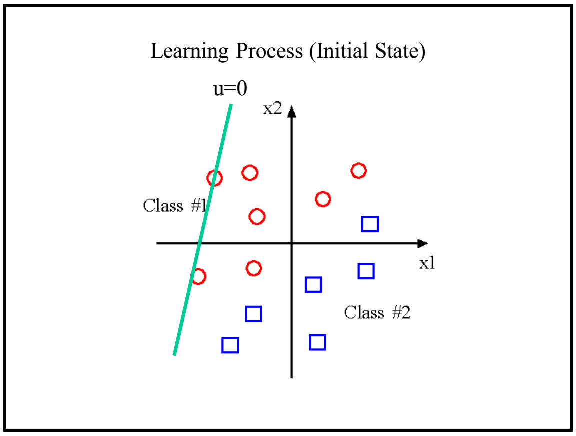


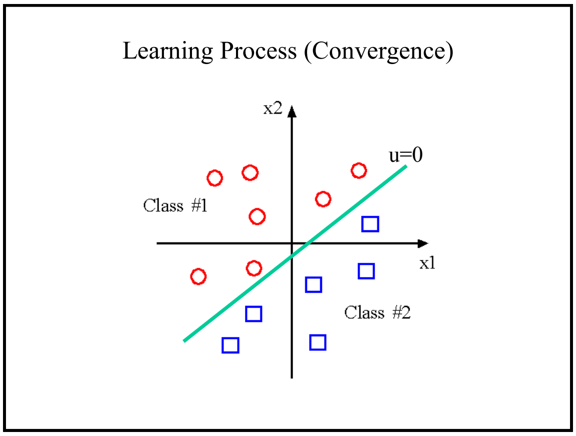
Learning of Connection Weights in Single-Layer NN

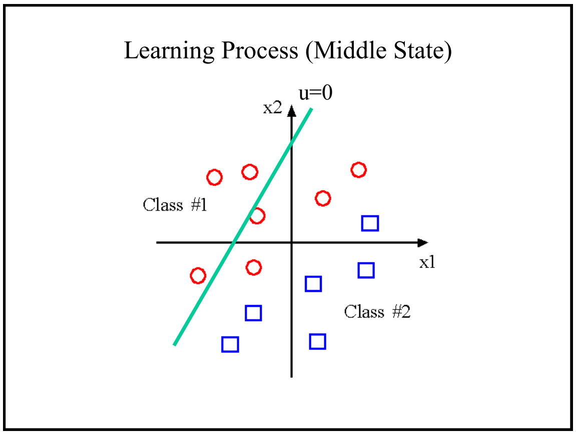


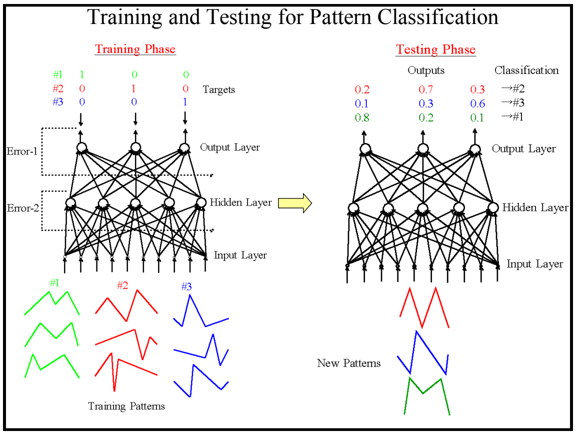

Learning of Connection Weights in Multi-Layer NN
- Error Back Propagation Algorithm -
d
y
Output
Gradient Method
δ j
wj
y1
y2
Chain Rule in
Derivative
δ i
x0=1
wji
Hidden
Input
x0=1
x1
x2


Learning Process (Middle State)
x2
u=0
Class #1
x1
Class #2


Learning Process (Initial State)
u=0
x2
Class #1
x1
Class #2


Learning Process (Convergence)
x2
u=0
Class #1
x1
Class #2


Learning Process (Middle State)
u=0
x2
Class #1
x1
Class #2


Training and Testing for Pattern Classification
Training Phase
#1 1 0 0
#2 0 1 0
#3 0 0 1
Targets
Error-1
Output Layer
Error-2
Hidden Layer
Input Layer
#1
#2
#3
Training Patterns
Testing Phase
Outputs
Classification
0.2 0.7 0.3 →#2
0.1 0.3 0.6 →#3
0.8 0.2 0.1 →#1
Output Layer
Hidden Layer
Input Layer
New Patterns

Application 1

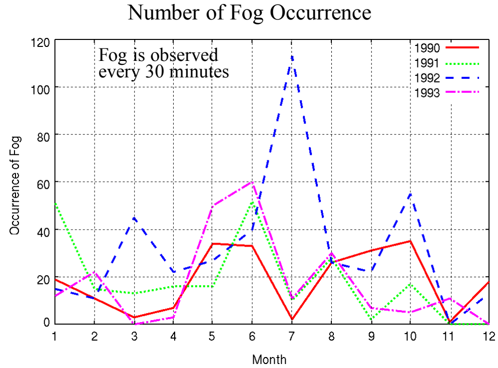
Prediction of Fog Occurrence

Weather Data

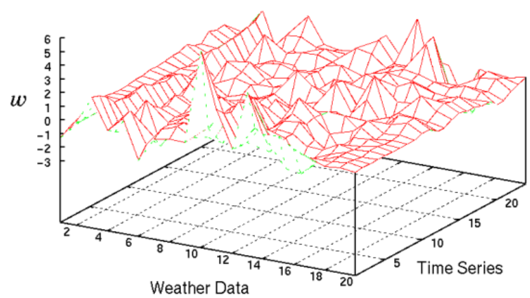
- Temperature
- Atmospheric Pressure
- Humidity
- Force of Wind
- Direction of Wind
- Cloud Condition
- Past Fog Occurrence
- .....

20 kinds of weather data are used

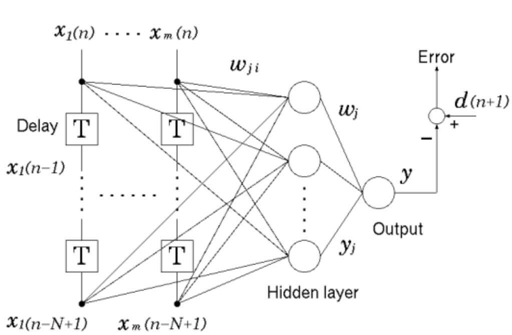
Number of Fog Occurrence



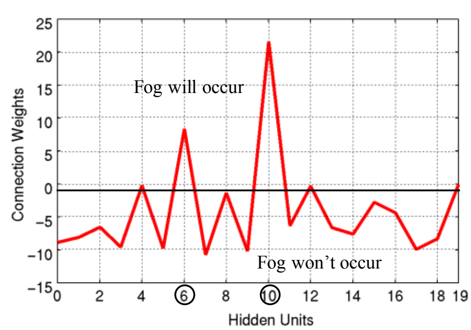
Connection Weights from Input to Hidden Unit



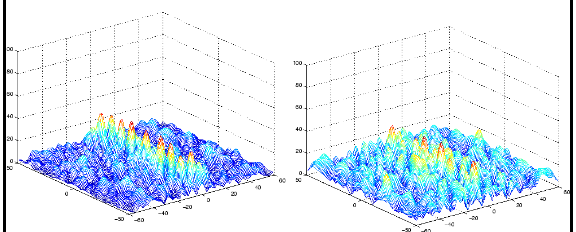
Neural Network for Prediction



Connection Weights from Hidden to Output



FFT of Connection Weights Used for Predicting Fog

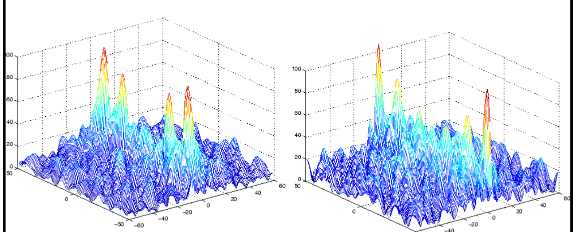


Input→Hidden Unit #6   Input→Hidden Unit #10

Application 2

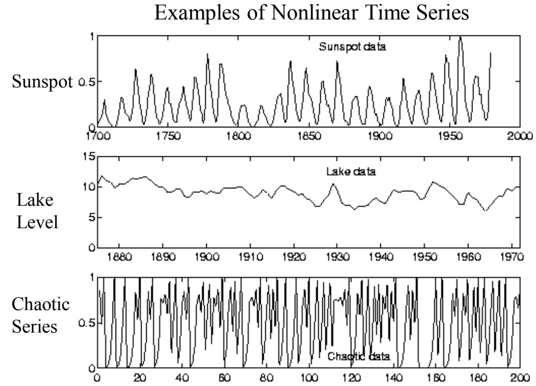
Nonlinear Time Series Prediction

FFT of Connection Weights for Predicting No Fog

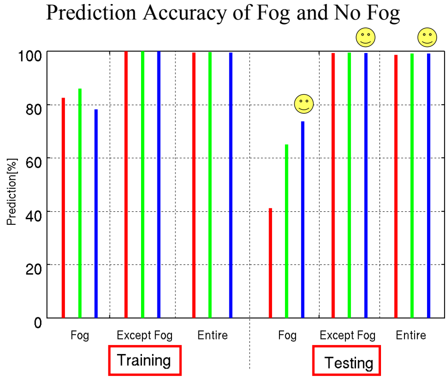


Input→Hidden Unit #3   Input→Hidden Unit #14

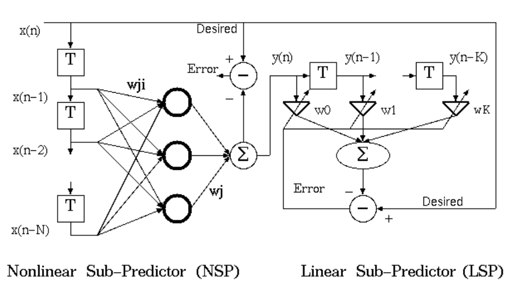
Examples of Nonlinear Time Series



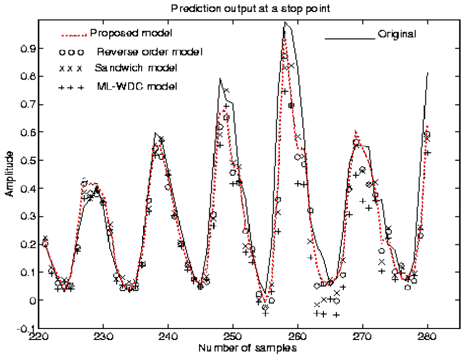
Prediction Accuracy of Fog and No Fog



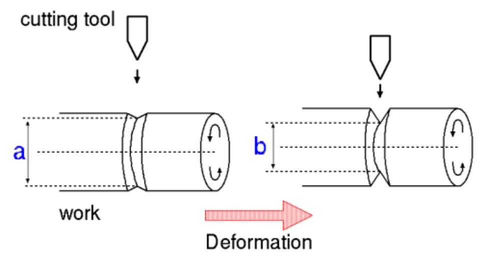
Nonlinear Predictor Combining NN and Linear Filter



Prediction Accuracy by Several Methods



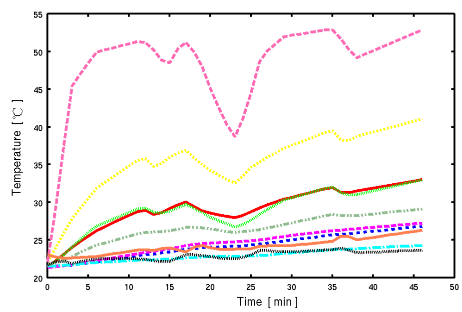
Deformation of Cutting by Temperature Change



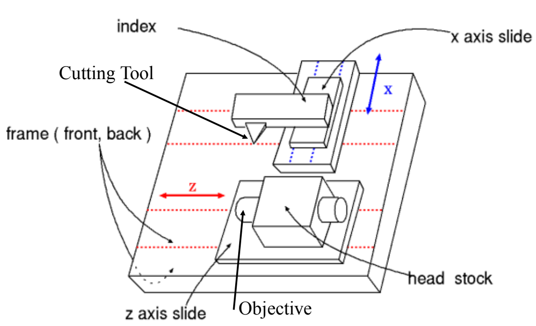
Application 3

Prediction of Machine Deformation

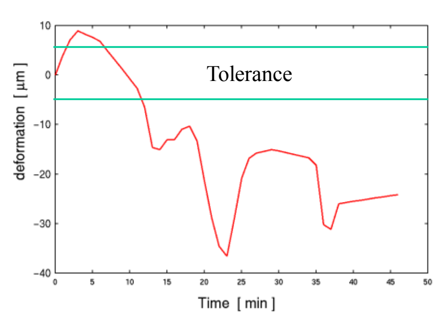
Machine Temperature Change in Time



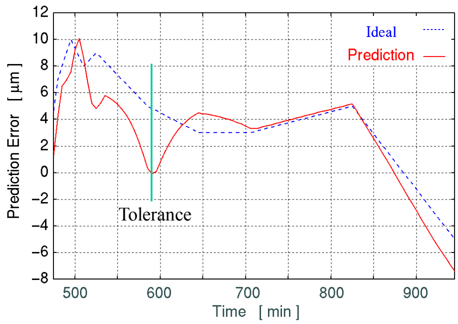
Numerically Controlled Cutting Machine



Deviation of Cutting by Temperature Change



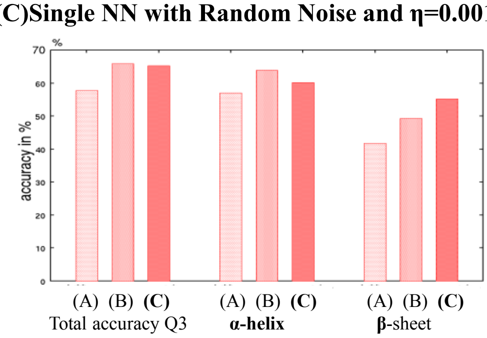
Prediction of Deformation Using NN



Application 4

Predicting Protein Secondary Structure

Comparison of Predition Accuracy in (A)Single NN in [6], (B)Single NN with η=0.00001, (C)Single NN with Random Noise and η=0.001

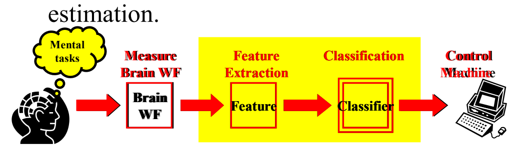


Application 5

Brain Computer Interface

Brain Computer Interface (BCI)

- Measure brain waveforms for subject thinking something (mental tasks).
- Analyze brain waveforms and estimate what kind of mental tasks does the subject imagine.
- Control computer or machine based on the estimation.



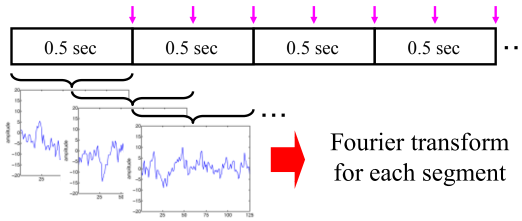
Approaches

- Feature
  Amplitude of Fourier transform
- Classification
  Multi-layer neural networks
- Brain waveforms
  Colorado State University
  http://www.cs.colorado.edu/eeg/

Five Mental Tasks

- Baseline: Nothing to do (Relax).
- Multiplication: Calculate 49 × 78 for example.
- Letter: Writing a sentence of letter.
- Rotation: Imagine rotating a 3-D object.
- Count: Writing numbers in order on a board.

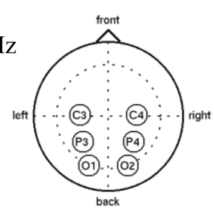
Segmental Processing



- Brain waveform of 10 sec is divided into segments of 0.5 sec.
- Mental tasks are estimated at each 0.25 sec.(↓)

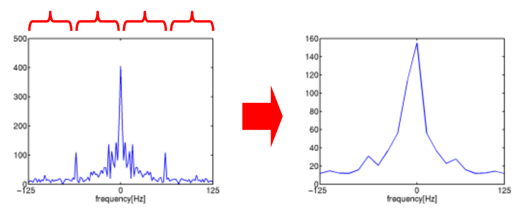
Measuring Brain Waveform

- Number of electrodes: 7ch
  C3, C4, P3, P4, O1, O2, EOG
- Measuring time: 10 sec
- Sampling frequency: 250Hz
  2500 samples per channel



Reduction of # of Samples

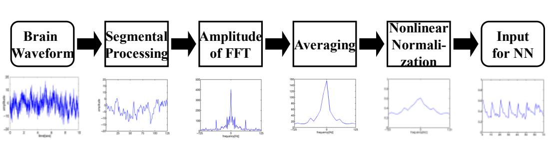
- # of samples are reduced from 125 to 20 by averaging the successive samples of waveform.



# of samples: 125   # of samples: 20

Pre-Processing of Brain Waveform

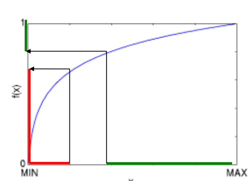
- Segmental processing
- Amplitude of Fourier transform
- Reduction of # of samples by averaging
- Nonlinear normalization of data



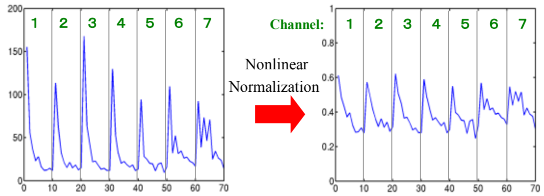
Nonlinear Normalization for Amplitude of Fourier Transform

- Amplitude of FFT is nonlinearly normalized in order to use samples having small values.

$$f(x) = \log(x - \min + 1) / \log(\max - \min + 1)$$



Nonlinear Normalization for Amplitude of Fourier Transform



7 channels are arranged at input nodes (10 × 7=70 samples)

Classification Accuracy for Subject 1 and 2

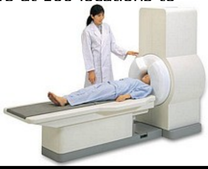
	Training Data			Test Data		
Subject	Correct	Error	Ratio	Correct	Error	Ratio
1	**99.7**	**0.1**	**0.99**	**79.7**	**10.5**	**0.88**
2	**95.5**	**0.8**	**0.99**	**45.5**	**33.7**	**0.57**

Simulation Setup

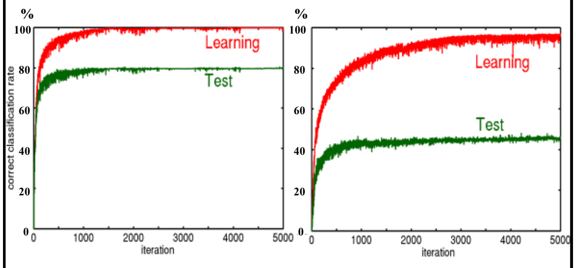
- 2 subjects
- Hidden units: 20
- Learning rate: 0.2
- Initial connection weights:
  Random numbers distributed during -0.2~0.2
- Threshold for rejection: 0.8

MEG (Magnetoencephalograph)

- A measurement instrument specifically designed to measure electrophysiological cerebral nerve activities.
- High time and spatial resolution performance
- SQUID fluxmeters, which detect the extremely weak magnetic field generated by the brain.
- MEGvision places the SQUID fluxmeters at 160 locations to cover the entire head.
- Complex magnetic field source generated by the activity of the brain can be recorded at a high spatial resolution.

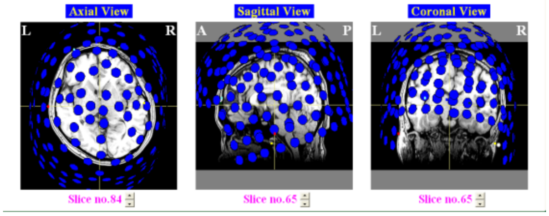


Learning Curves for Training and Testing Data Sets

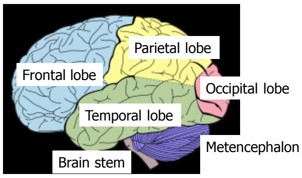


Subject 1 Subject 2

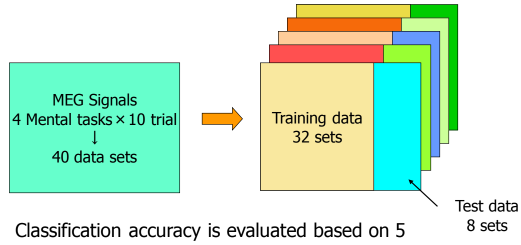
Layout of Sensors on Head



Channel (Sensor) Selection



Performance Evaluation



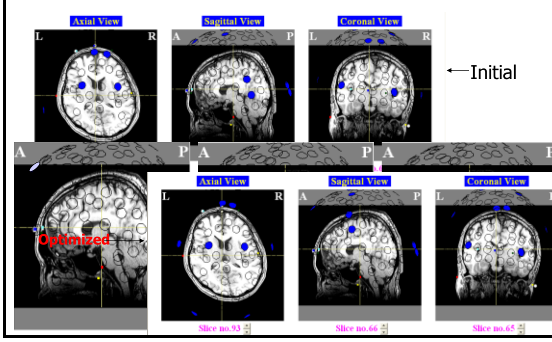
Classification accuracy is evaluated based on 5 kinds of combinations and their average.

Channel (Sensor) Selection

8 channels are selected from 8 main lobes. The initial location is set to the central point of each lobes.

Ch1: Frontal lobe (left), Ch2: Frontal lobe (right)
Ch3: Parietal lobe (left), Ch4: Parietal lobe (right)
Ch5: Temporal lobe (left), Ch6: Temporal lobe (right)
Ch7: Occipital lobe (left), Ch8: Occipital lobe (right)

Optimization of Sensor Location



Mental Tasks

Four kinds of mental tasks are used.

- Baseline: Staying in relaxed condition
- Multiplication: a 3-digit number by a 1-digit number (ex. 456 × 8)
- Sports: Playing some sport, which is determined by the subject.
- Rotation: Rotating some object, which is determined by the subject.

Classification Rates

Correct/Error

	Subject 1	Subject 2	Subject 3
Sensor Location Initial [%]	**90.0/10.0**	**82.5/17.5**	**57.5/42.5**
Sensor Location Optimized [%]	**97.5/2.5**	**85.0/15.0**	**72.5/27.5**

Classification Score (Subject 1)

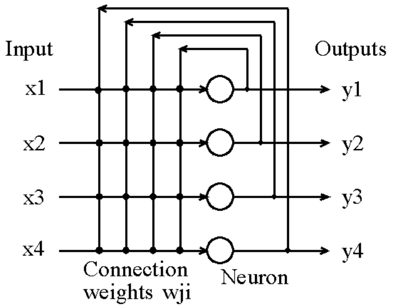
Mental tasks	B	M	S	R	Correct [%]	Error [%]
B	10	0	0	0	100	0
M	0	10	0	0	100	0
S	1	0	9	0	90	10
R	0	0	0	10	100	0
				Av.	97.5	2.5

Recurrent Neural Networks

Classification Score (Subject 2)

Mental tasks	B	M	S	R	Correct [%]	Error [%]
B	9	1	0	0	90	10
M	1	9	0	0	90	10
S	1	1	7	1	70	30
R	0	0	1	9	90	10
				Av.	85.0	15.0

Recurrent Neural Network



Classification Score (Subject 3)

Mental tasks	B	M	S	R	Correct [%]	Error [%]
B	4	4	1	1	40	60
M	1	8	1	0	80	20
S	1	0	8	1	80	20
R	1	0	0	9	90	10
				Av.	72.5	27.5

Hopfield Neural Network

- Symmetrical Connections $w_{ij} = w_{ji}$
- No Self-loop
- One neuron randomly selected is updated.
- The energy function always decrease or stay at the same value.
- Memory Capacity is about 15% of Neurons

Associative Memory (1)

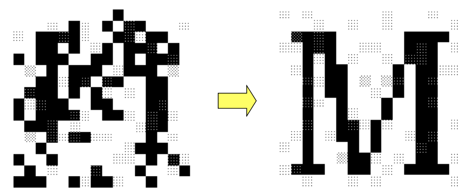
4x4=16 Neuron RNN
6 Random Patterns $\{p_i\}$ are Stored

Connection Weights

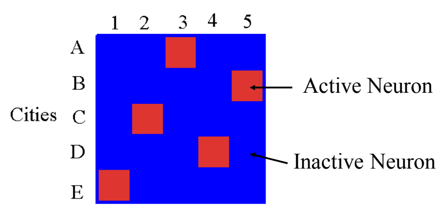
$$W=\sum_{i=1}^{M} p_i p_i^T$$

★Demonstration
Association from another random patterns

Association of 'M' from Its Noisy Pattern

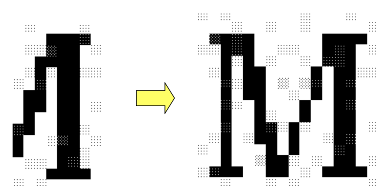


Traveling Salesman Problem



Network State (5 × 5 Neurons)
Traveling Schedule
E→C→A→D→B→E

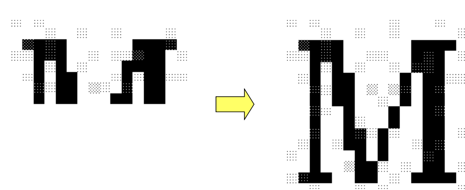
Association of 'M' from Its Right Half Pattern



Associative Memory (2)

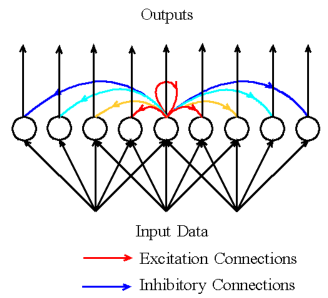
- Error Correction Learning with Hysteresis
- Adaptive Hysteresis Threshold for Association
- 51 Alphabet Letters and 10 Digits are Stored in 16x16=256 Neuron RNN. 25% of Neurons

Association of 'M' from Its Upper Half Pattern



Competitive Learning

Lateral Inhibition Model



END OF THIS LECTURE

THANK YOU