

センサーデータのためのニューラルネットワーク ～ 時系列データの分類と異常検知 ～

MathWorks Japan

アプリケーション エンジニアリング部 テクニカルコンピューティング
太田 英司

Agenda

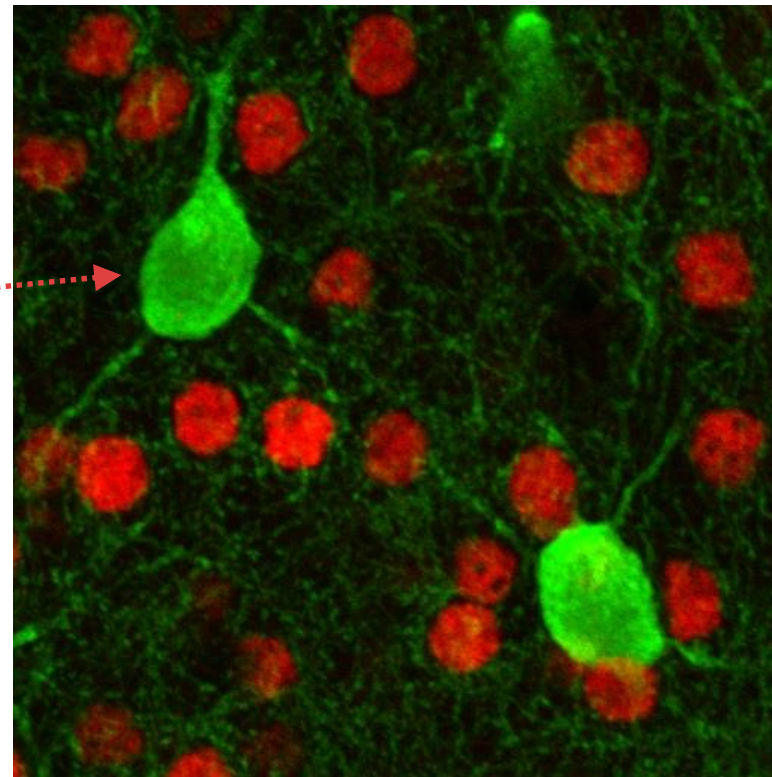
- ニューラルネットの基礎
- 時系列データの分類
 - 部分時系列
 - 自己符号化器 / 積層自己符号化器
 - LSTM (Long Short Term Memory)
- 異常検知 (自己符号化器)
 - 心電図の例題
 - 構造ヘルスマモニタリングの例題
- システムへの展開・Cコード生成
 - 学習済みネットワークのブロック化
 - MATLAB Production Server™

ニューラルネットワークとは？

神経細胞（ニューロン）の数学的なモデル化に起源を持つ学習器

神経細胞（ニューロン）

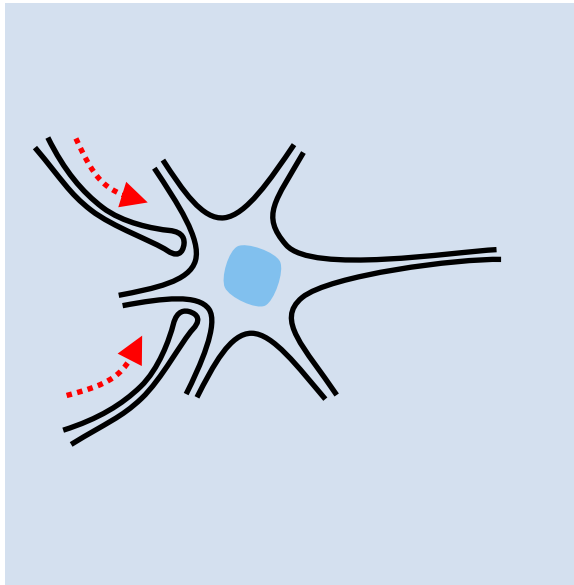
- 軸索によりネットワークを構成
- 電気的な興奮状態を伝え合う



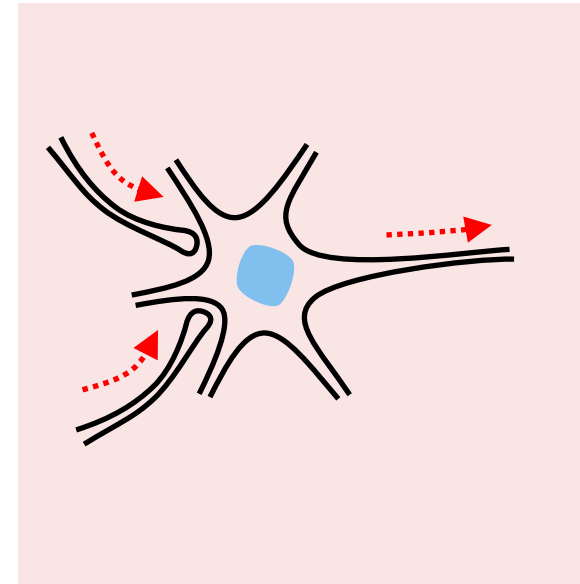
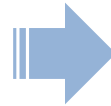
画像提供： 独立行政法人 理化学研究所様

ニューラルネットワークとは？

他の神経細胞からの興奮状態が伝わり、電位のレベルが上昇する



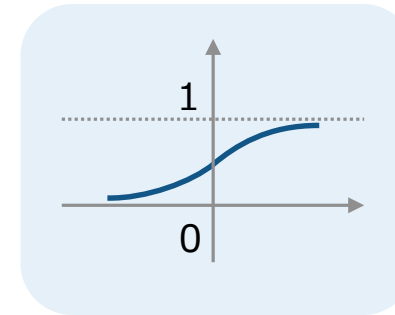
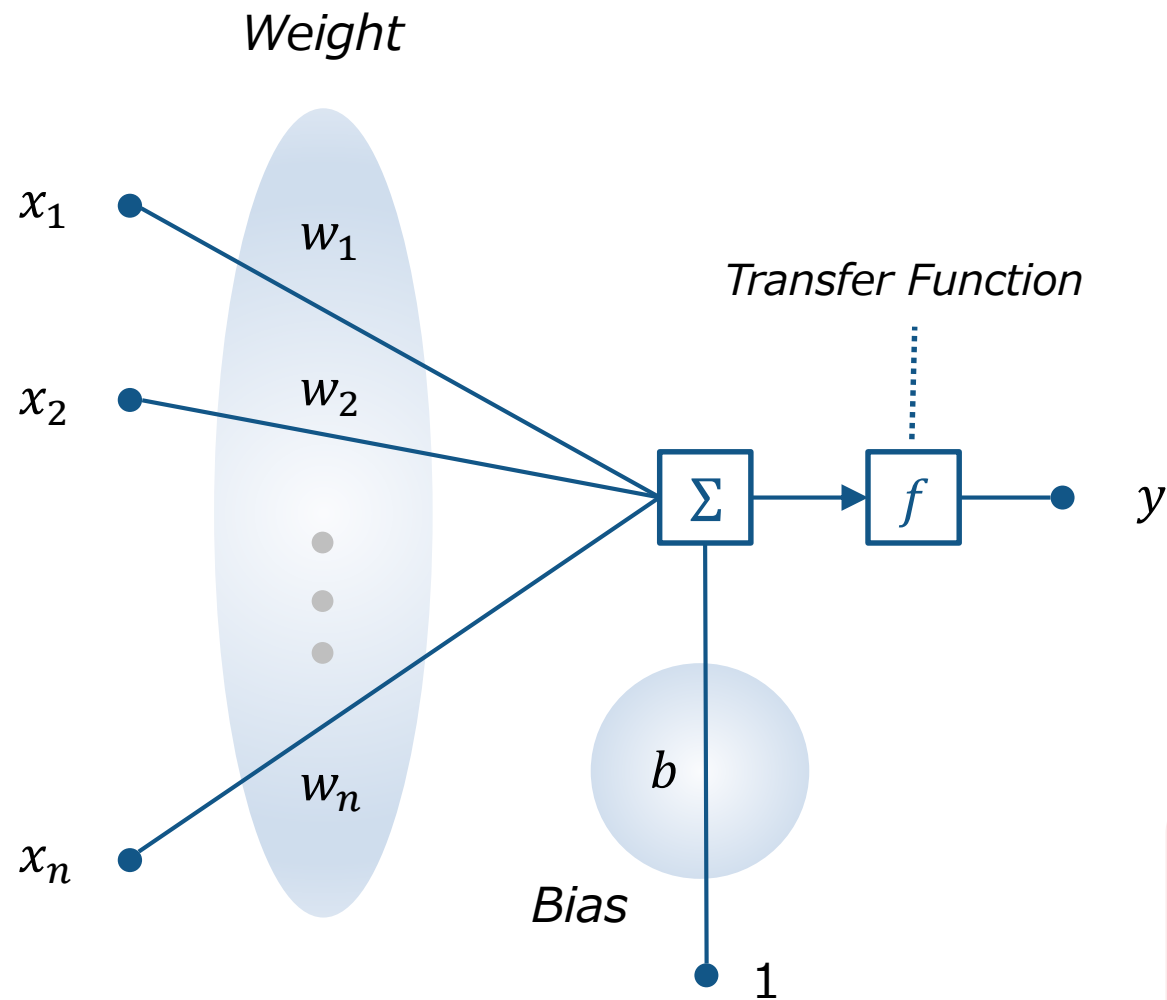
非興奮状態



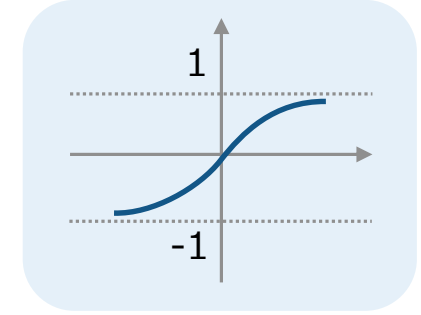
興奮状態

電位のレベルが閾値を超えると自身が発火する

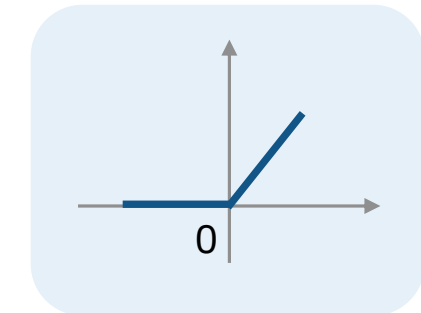
ニューラルネットワークとは？



Logistic Sigmoid



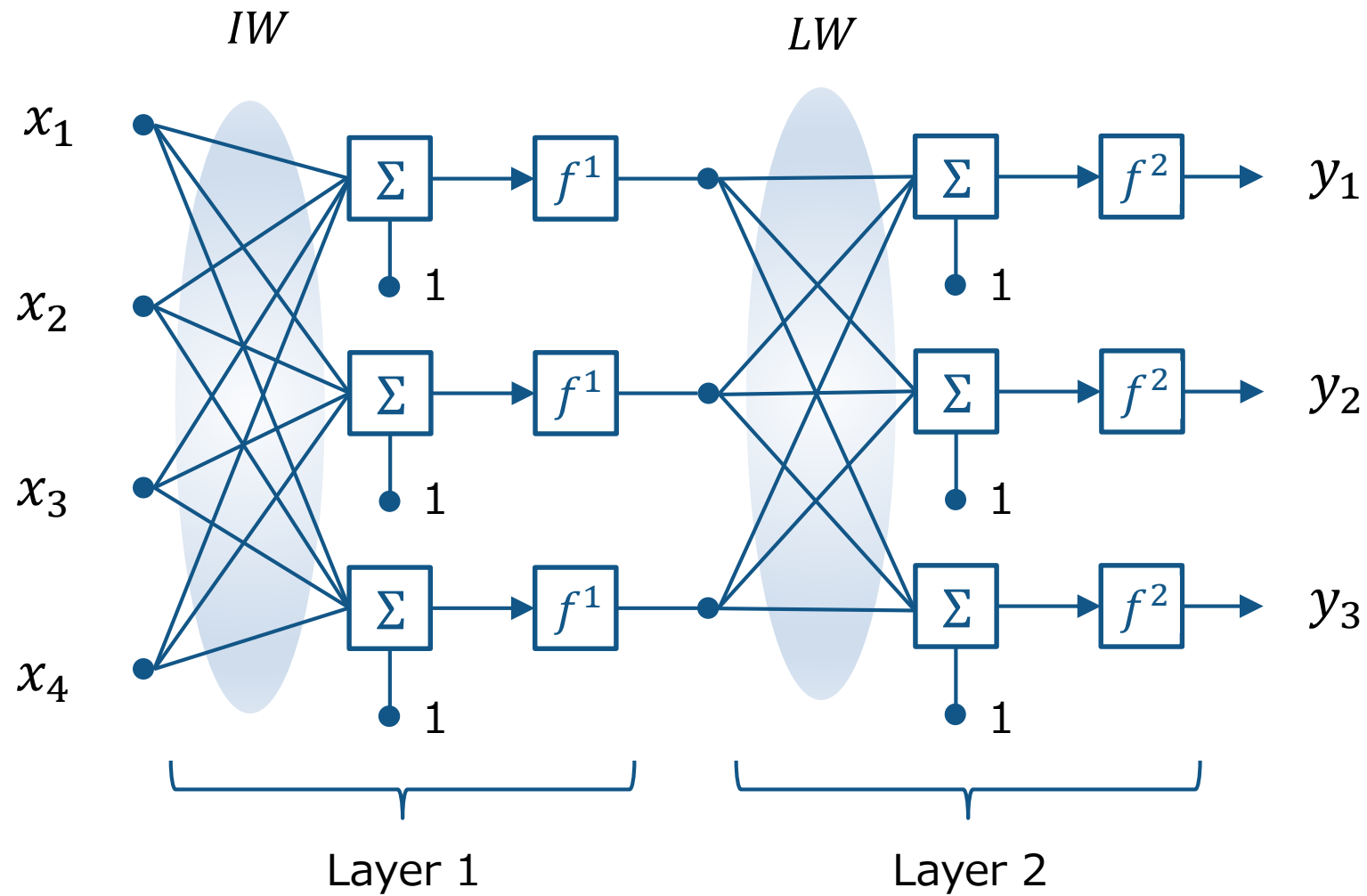
Tangent Sigmoid



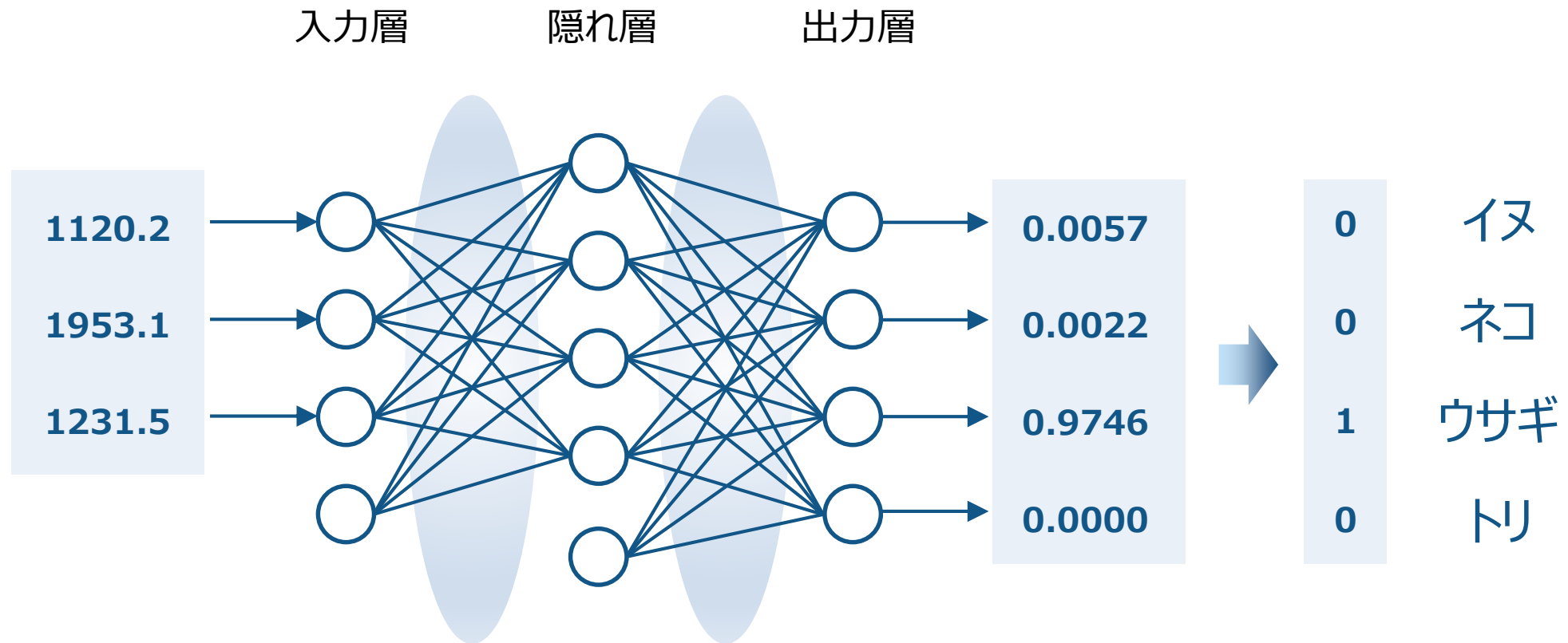
Rectified Linear Unit

$$y = f\left(\sum_{k=1}^n w_k \cdot x_k + b\right)$$

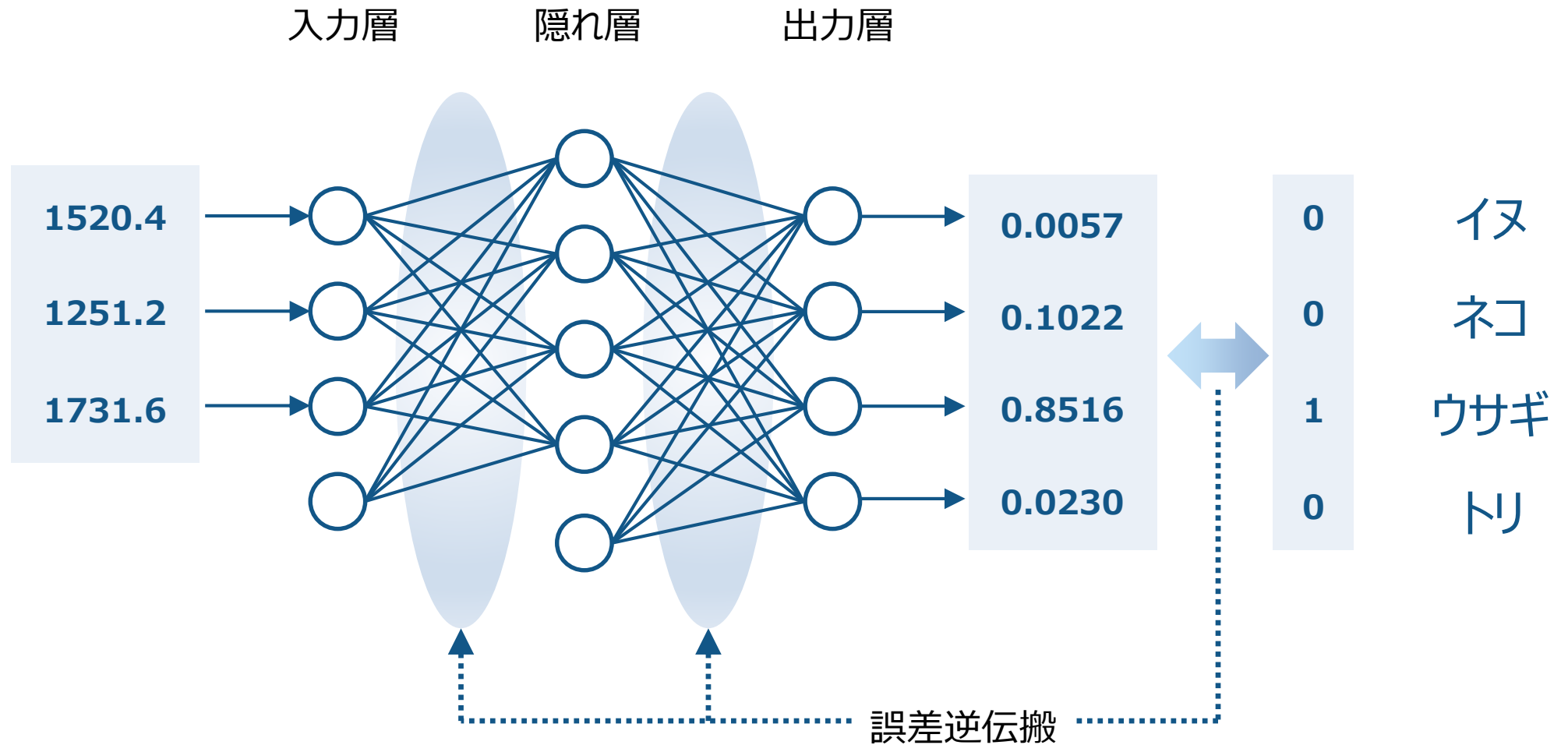
ニューラルネットワークとは？



ニューラルネットワークによる推論

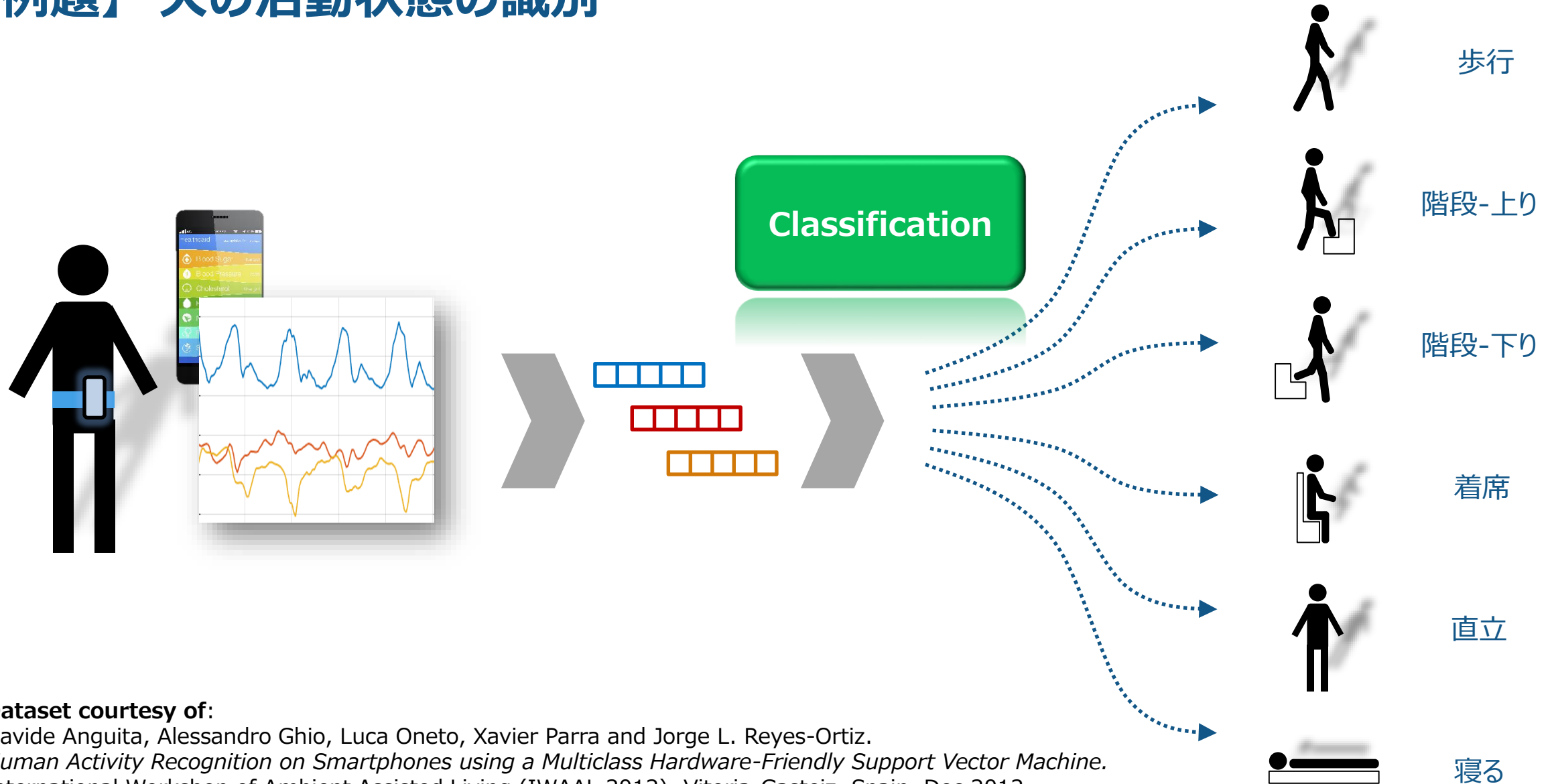


ニューラルネットワークの学習



時系列データの分類

【例題】人の活動状態の識別



Dataset courtesy of:

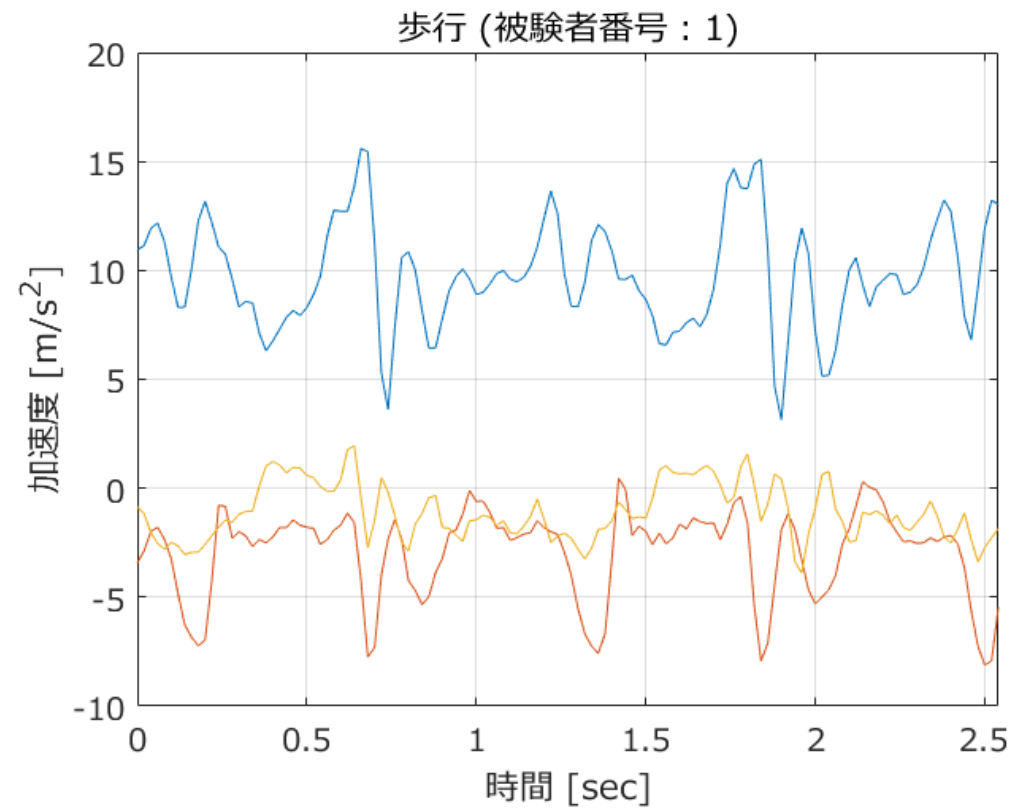
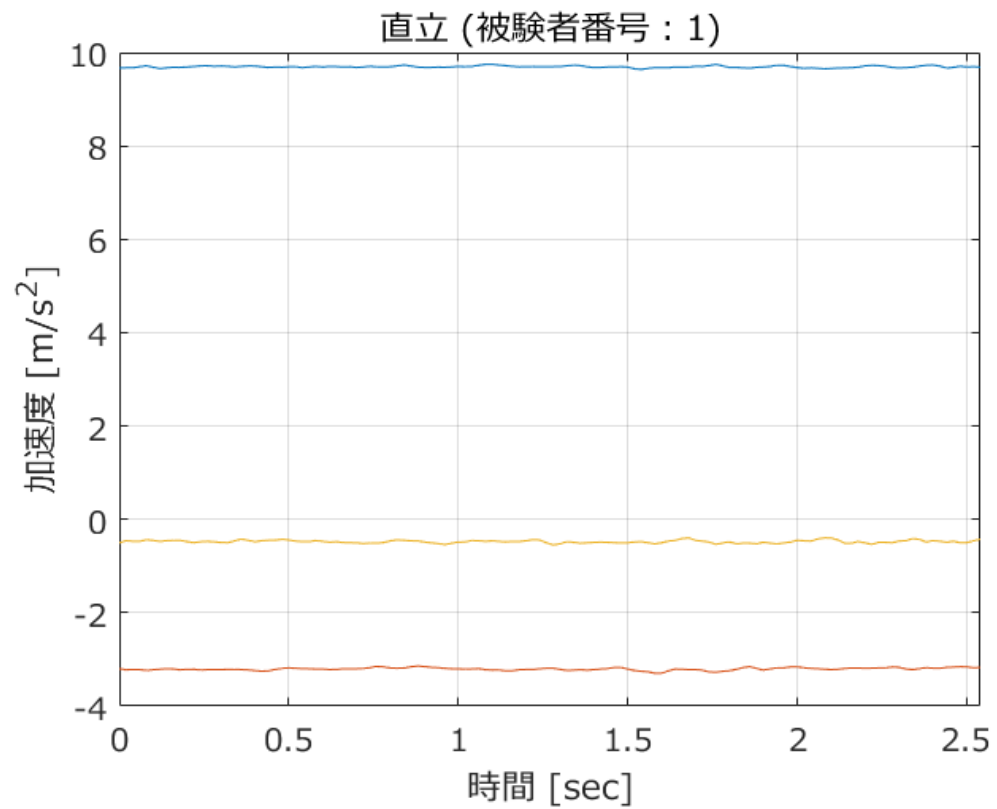
Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz.

Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine.

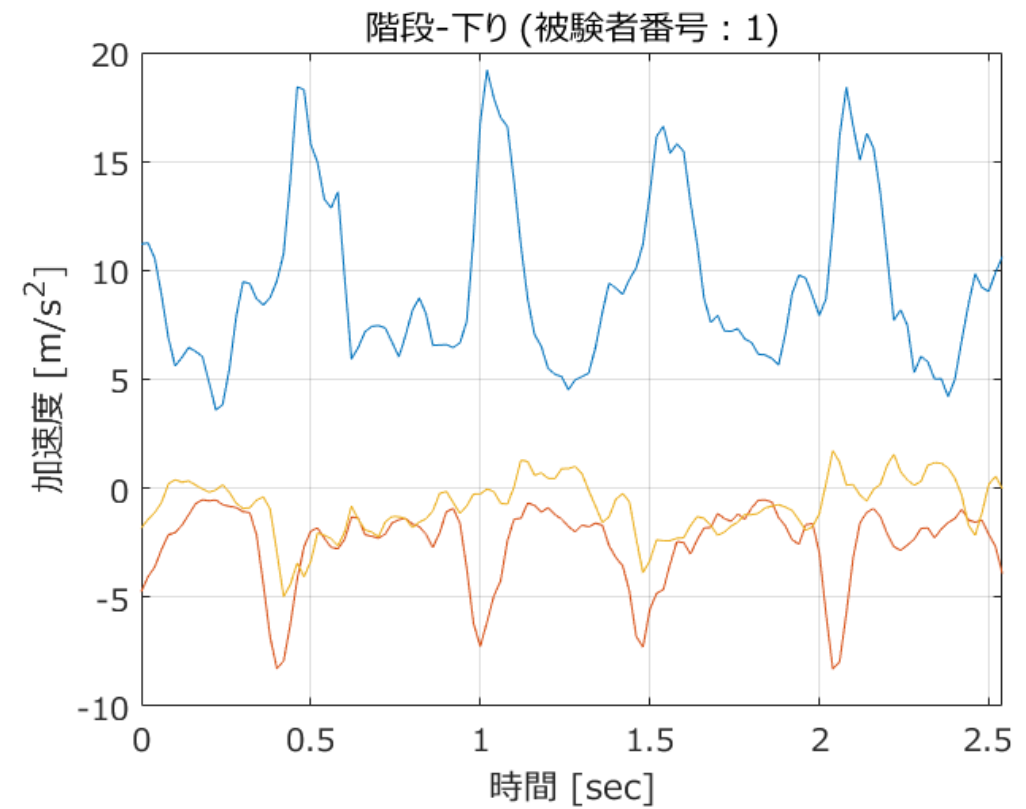
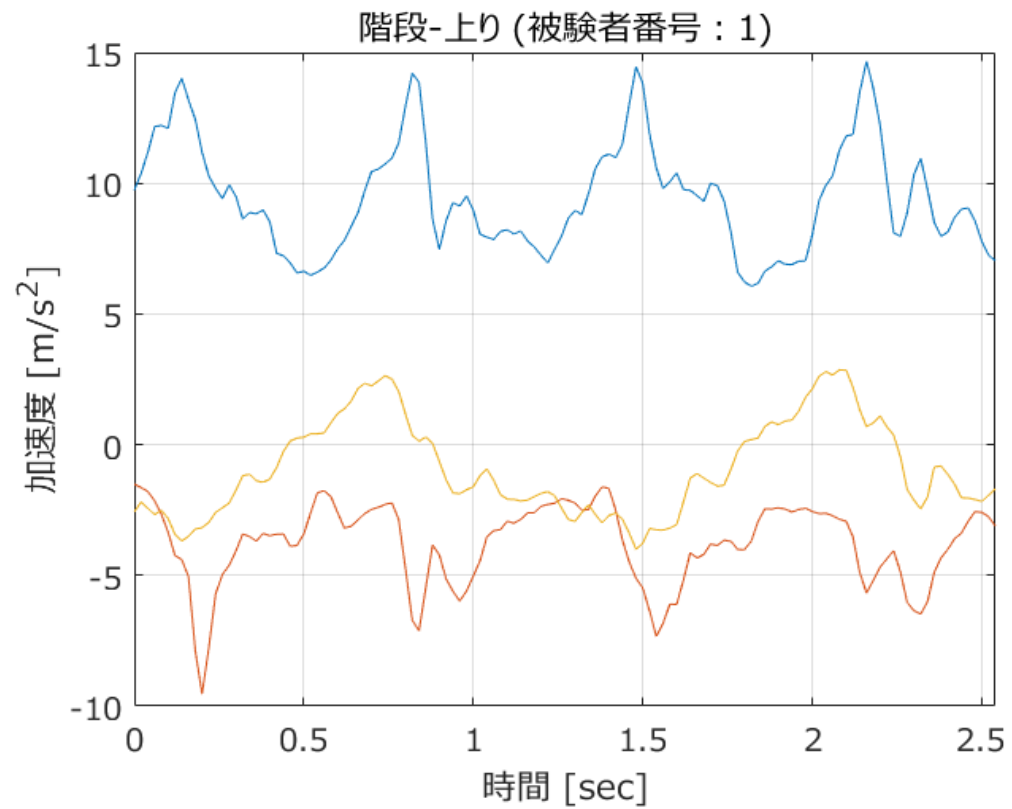
International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

【例題】 人の活動状態の識別

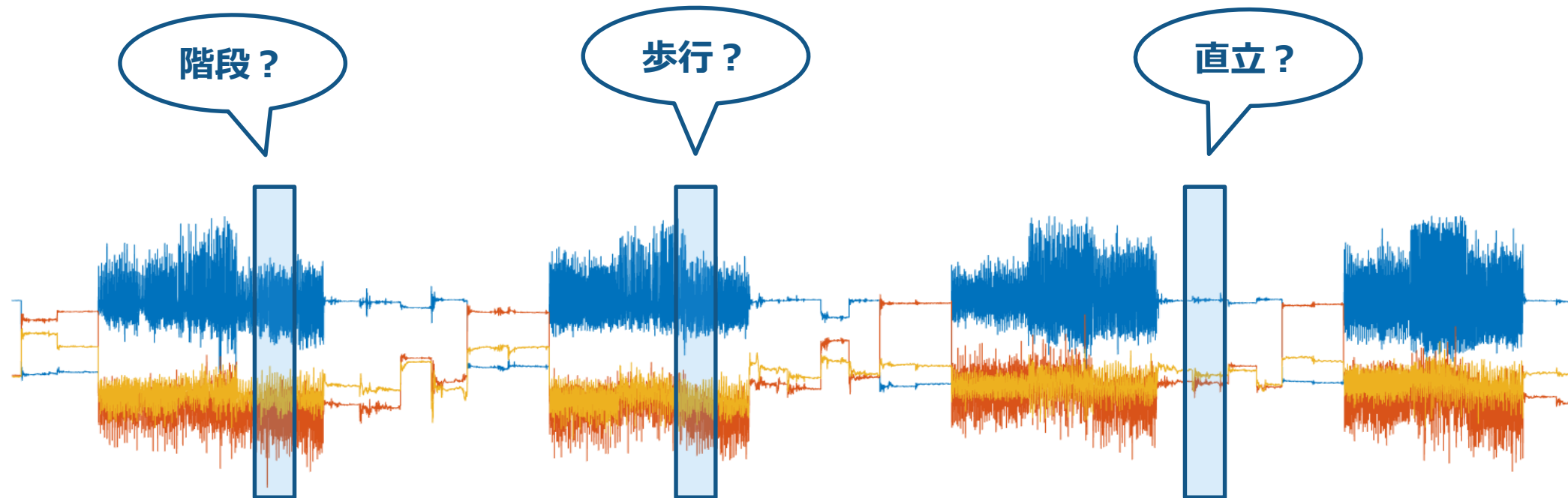


【例題】 人の活動状態の識別



【例題】 人の活動状態の識別

3 軸の加速度センサーの情報を使って、人の活動状態を識別したい

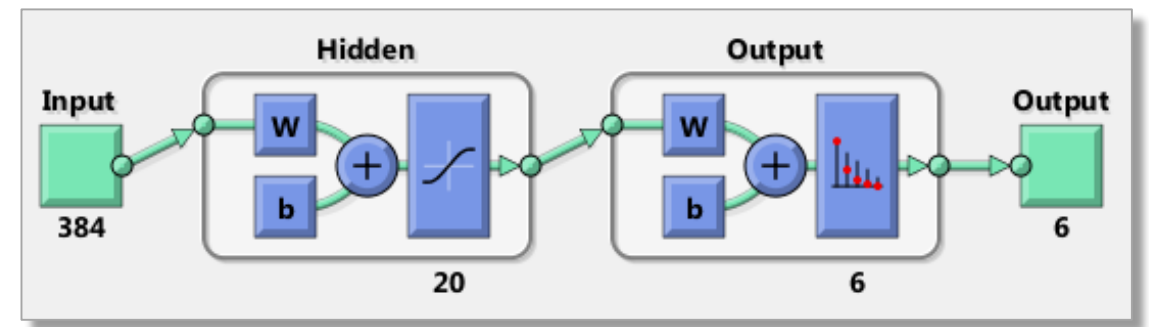
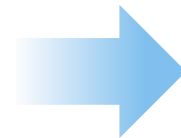
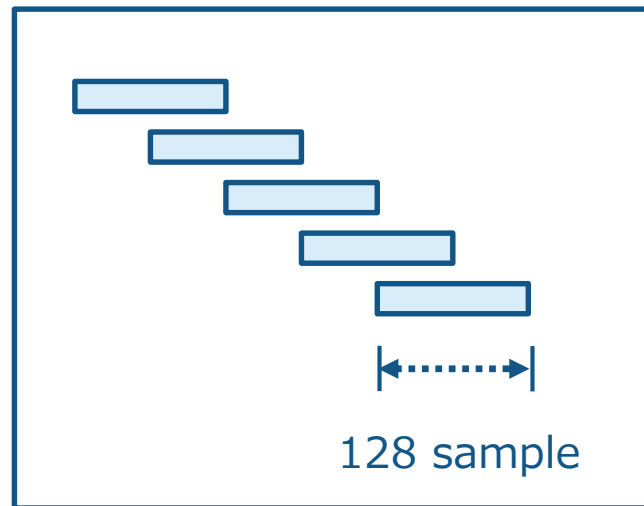


【例題】 人の活動状態の識別

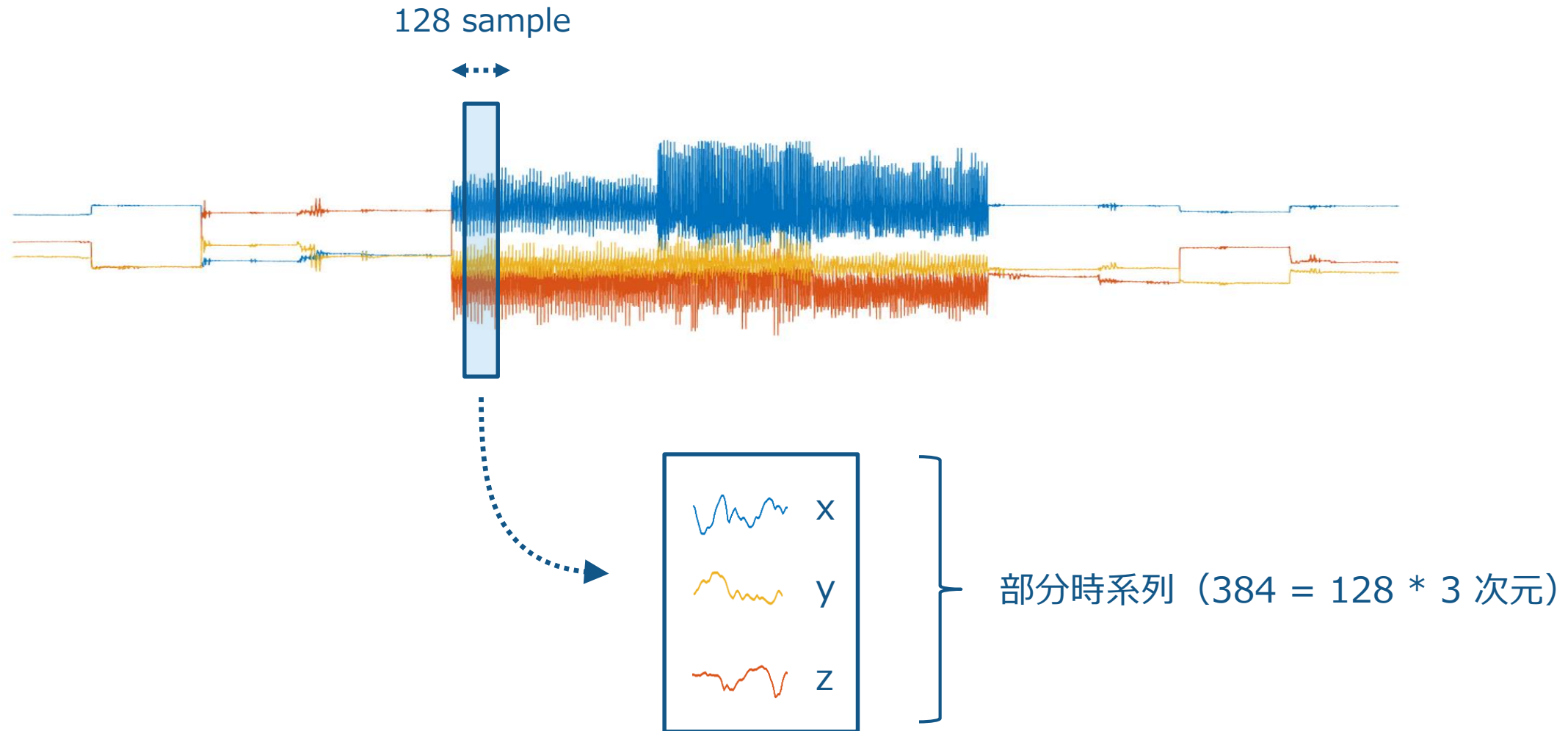
3軸加速度センサーのデータを窓幅128で切り出し、ニューラルネットより活動状態を識別する



部分時系列（窓幅 128）

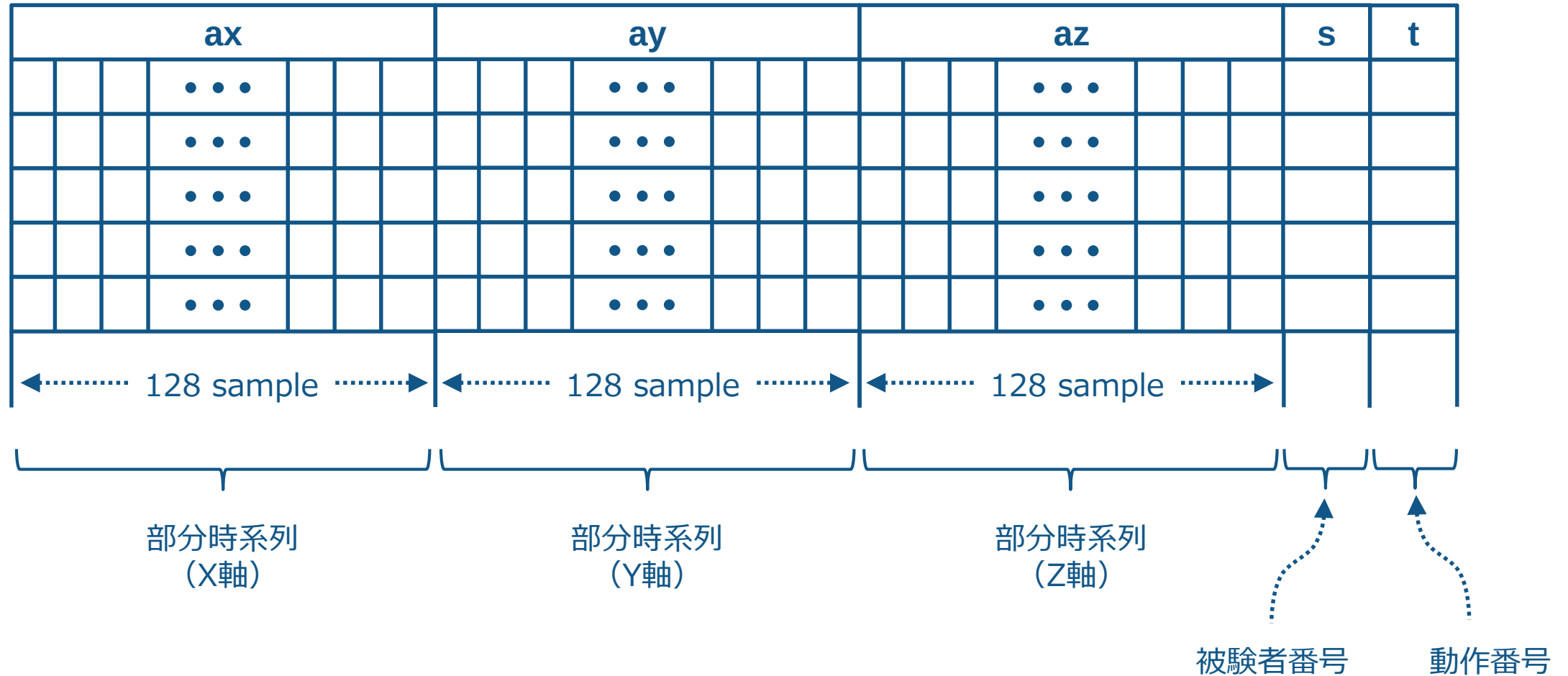


【例題】 人の活動状態の識別



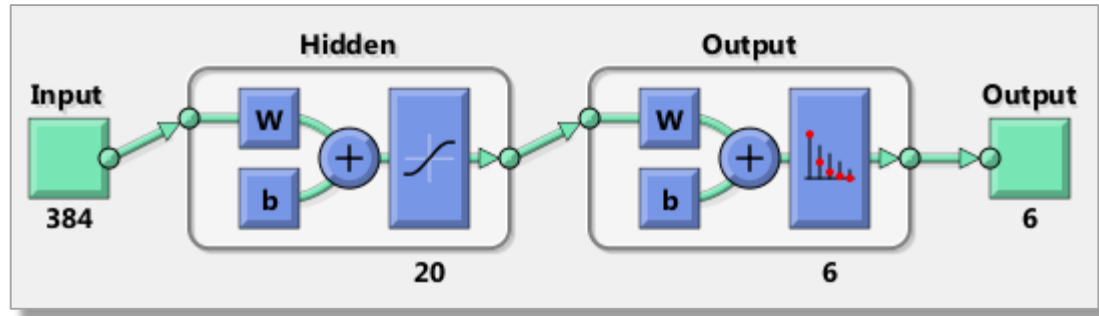
【例題】 人の活動状態の識別

加速度センサーの情報は一定の窓幅で部分時系列としてテーブルに格納されている

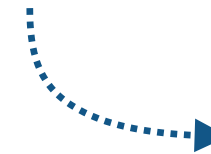


【例題】 人の活動状態の識別

3 軸の加速度センサーのデータを使って、人の活動状態を識別したい



入力 : 384 次元
部分時系列 (加速度)



⋮	⋮	⋮	⋮	⋮

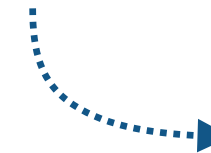
```
net = patternnet(20);
```

```
net = train(net, x, t);
```

```
t_hat = net(x);
```

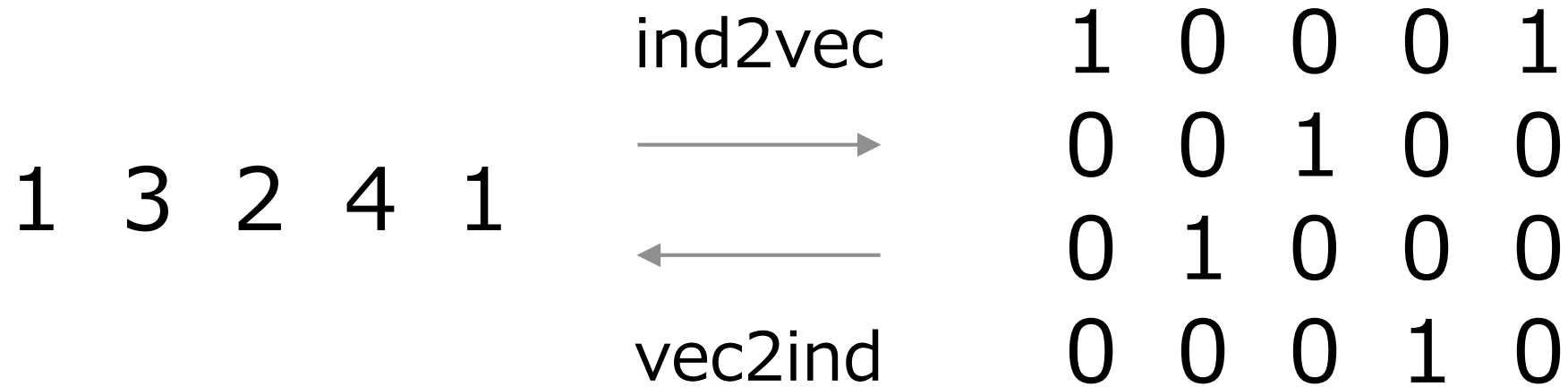
```
plotconfusion(t, t_hat)
```

出力 : 6 次元
人の活動状態 (ラベル)



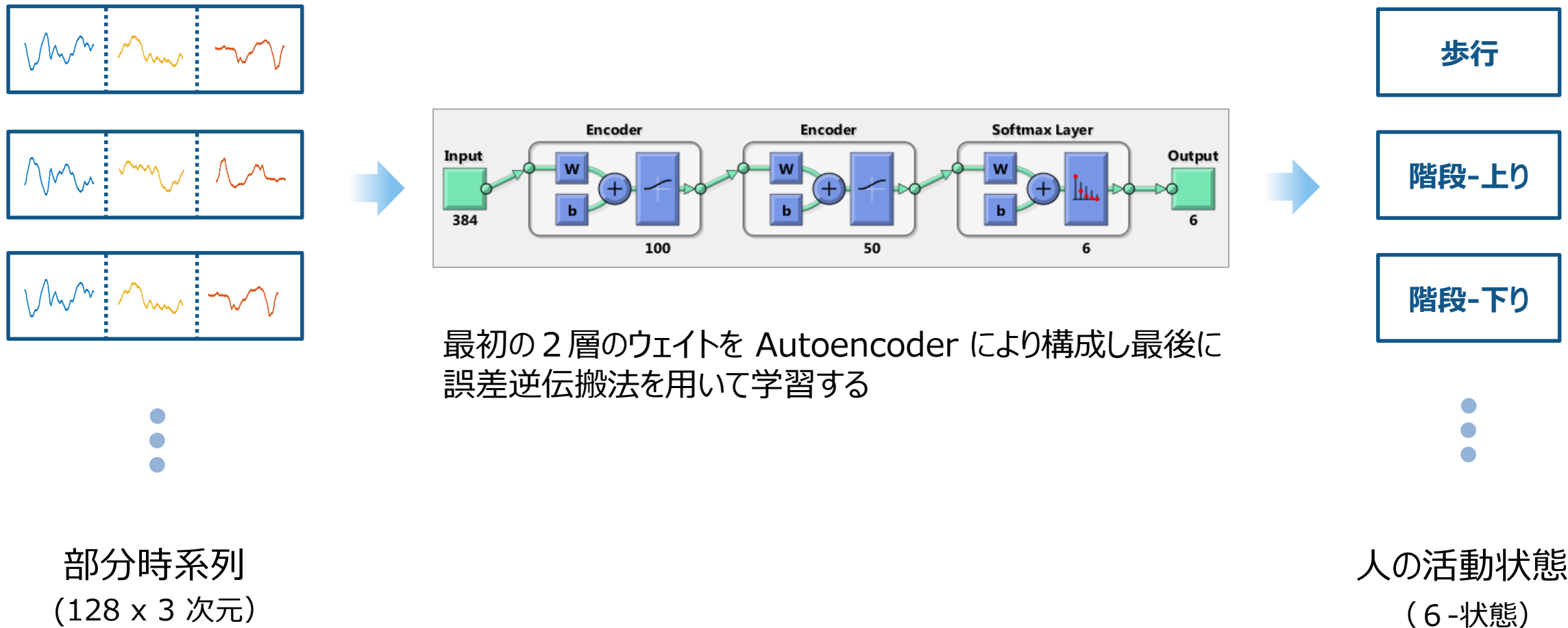
1	0	0	0	0
0	0	1	0	0
0	1	0	0	0
0	0	0	0	0
0	0	0	1	1
0	0	0	0	0

1 of K 符号化 (one-hot ベクトル表現)



【例題】人の活動状態の識別

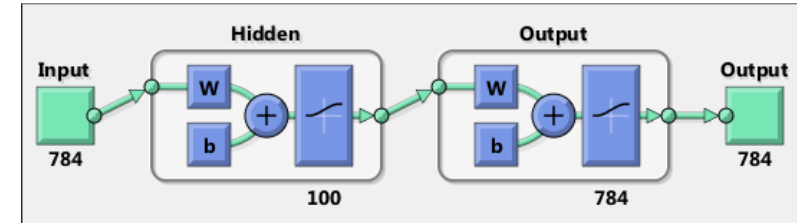
Stacked Autoencoder による人の活動状態の識別



自己符号化器 (Autoencoder) とは？

入力したデータと同じものが出力側で再生されるように学習させるネットワーク

- 恒等写像となるようなウェイトを学習
- 主に DNN の事前学習に利用
- 主成分分析・スパース符号化とも関連

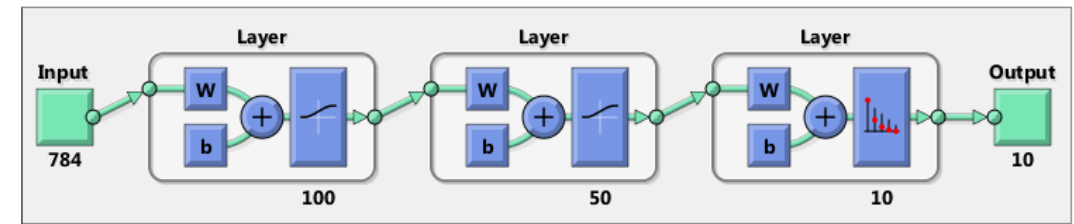


自己符号化器 (Autoencoder)

【コード例】 自己符号化器の学習

```
autoenc = trainAutoencoder(X, hiddenSize,...  
    'L2WeightRegularization', 0.001, ...  
    'SparsityRegularization', 4, ...  
    'SparsityProportion', 0.05, ...  
    'DecoderTransferFunction', 'purelin'  
    'useGPU', true);
```

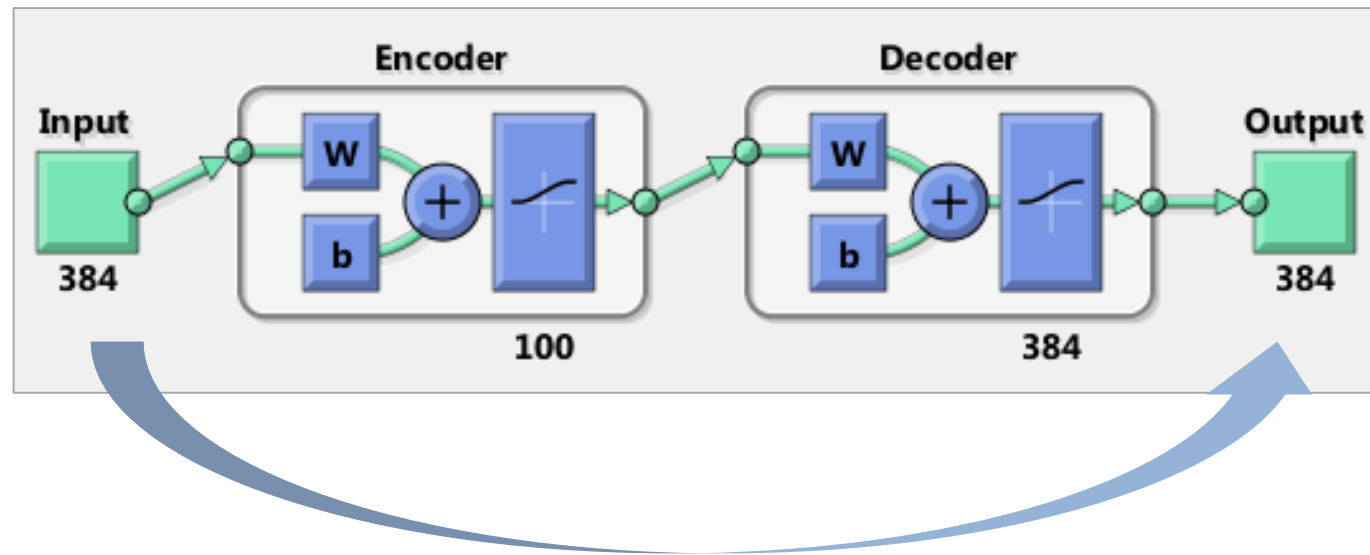
←..... GPUによる高速化



積層自己符号化器 (Stacked Autoencoder)

Step 1 : Autoencoder によるウェイトの学習

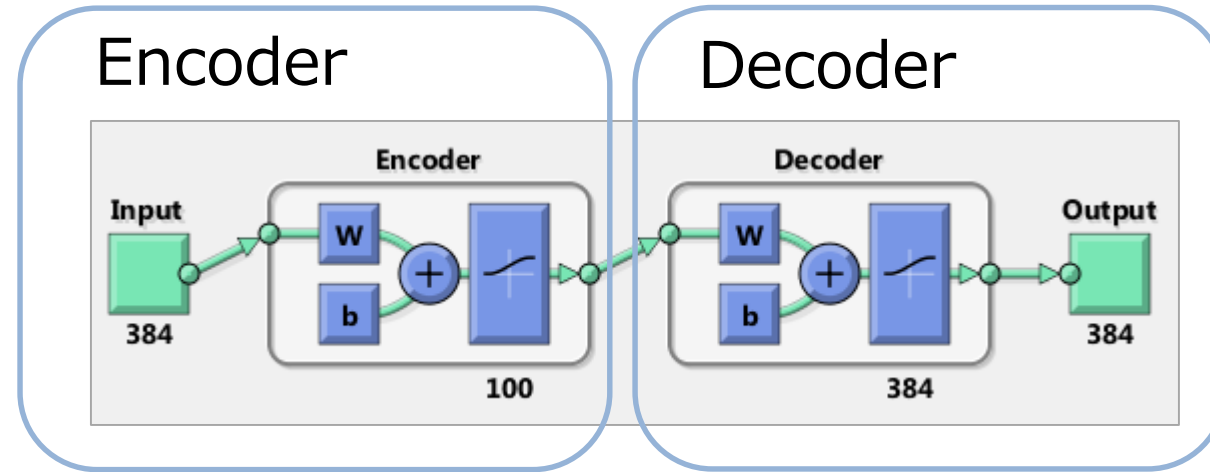
Autoencoder



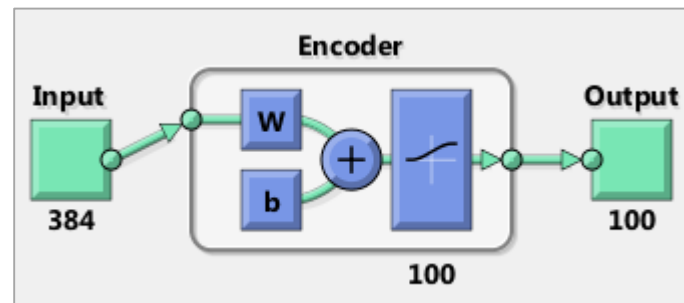
上のネットワークが恒等写像になるようにウェイトを学習する

※ 正確には、MATLAB の Autoencoder は Sparse Autoencoder と呼ばれる中間層の活性に疎性を仮定したものになっている。詳細は後述の PDF を参照

Step 2 : Encoder 部分のみ残す

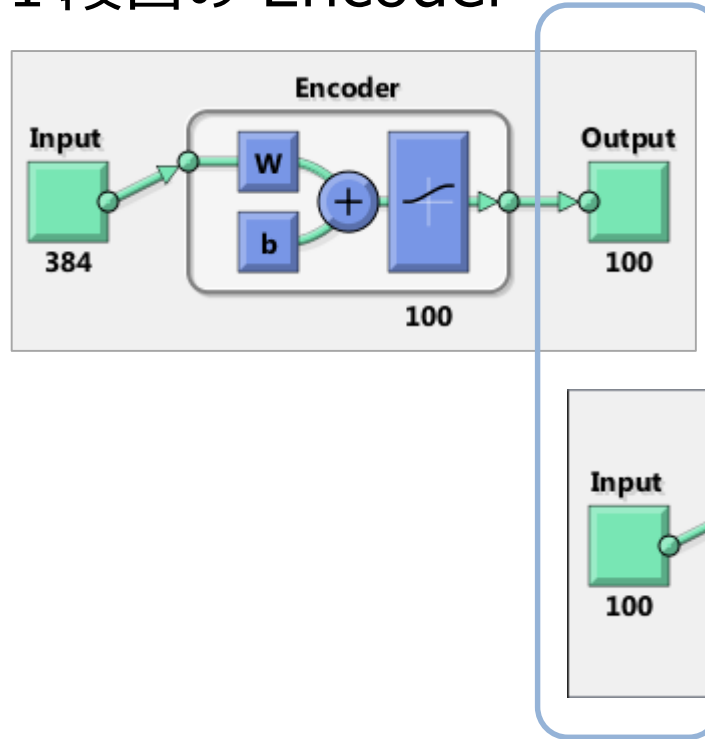


Encoder 部分のみ残す

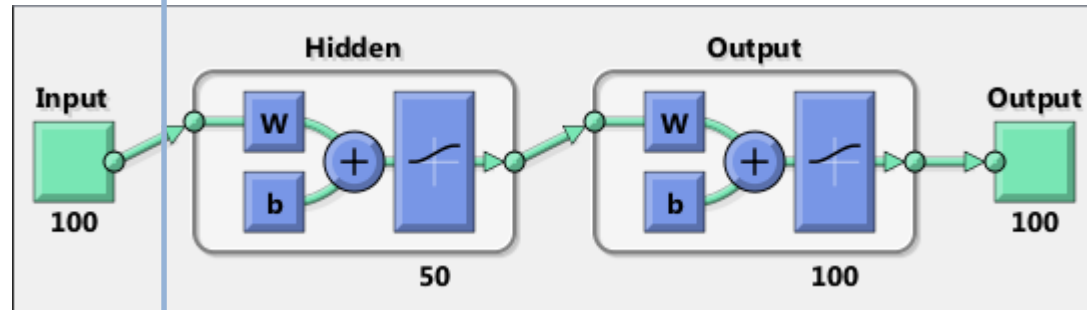


Step 3 : Autoencoder によるウェイトの学習

1 段目の Encoder



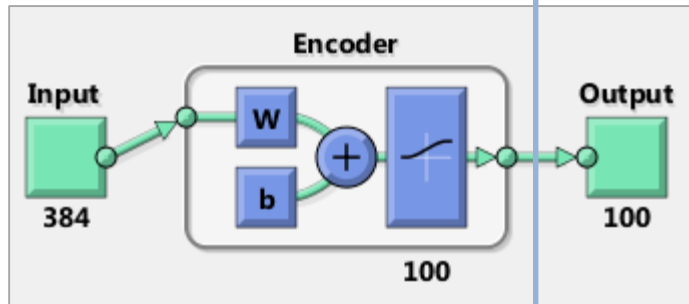
2 段目の Autoencoder



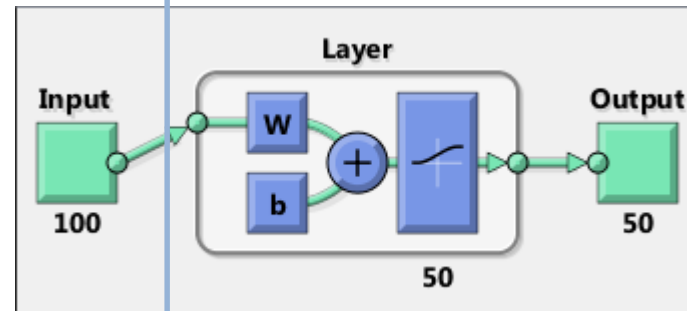
1 段目の Encoder 出力を使い、次の Autoencoder を学習させる

Step 4 : Encoder 部分のみ残す

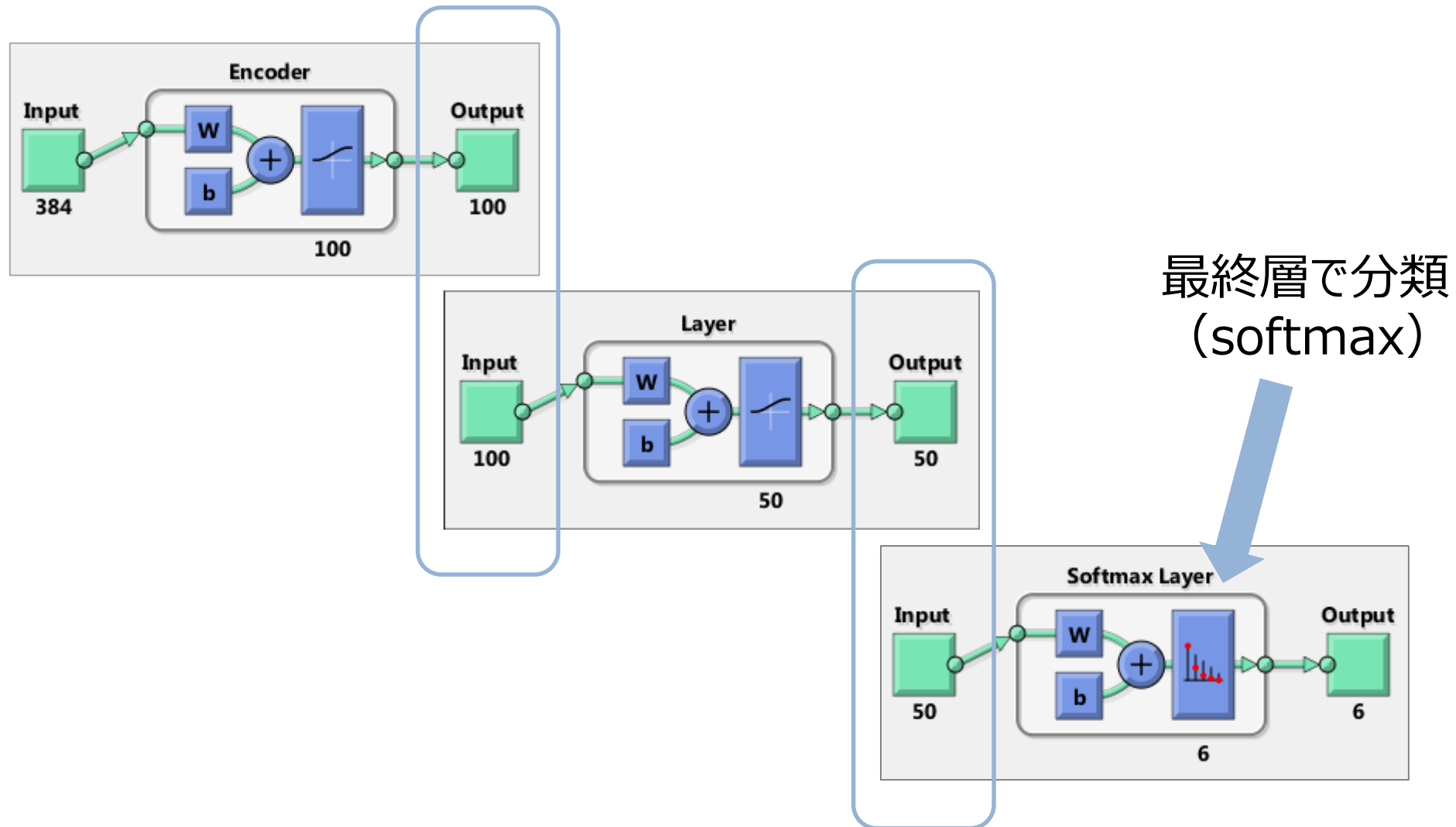
1 段目の Encoder



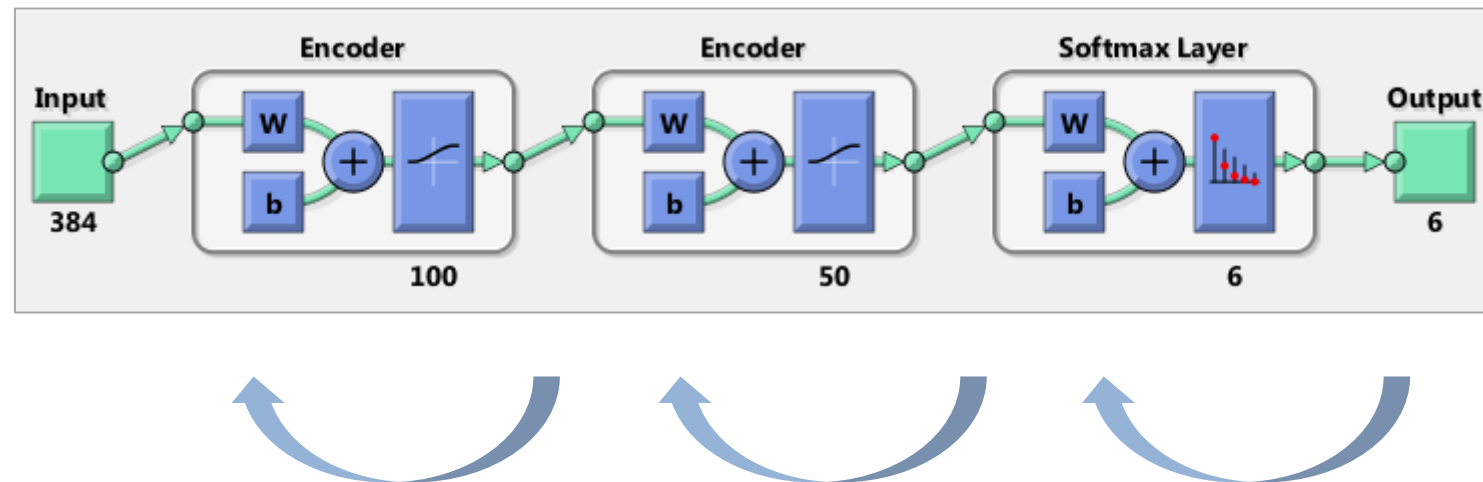
2 段目の Encoder



Step 5 : 分類のために最終層を追加



Step 6 : 微調整 (Fine Tuning) を行う



誤差逆伝搬法により、ウェイトの微調整 (Fine Tuning) を行う

【例題】人の活動状態の識別

積層自己符号化器として層を増やしたところ
分類性能が 3% 程度改善！

Accuracy : 0.81371

歩行	349	110	28	3	6	0
階段-上り	66	355	41	8	1	0
階段-下り	9	61	339	8	3	0
着席	1	23	0	360	107	0
直立	0	1	0	73	458	0
寝る	0	0	0	0	0	537
	歩行	階段-上り	階段-下り	着席	直立	寝る

混合行列（2層のネットワークの場合）

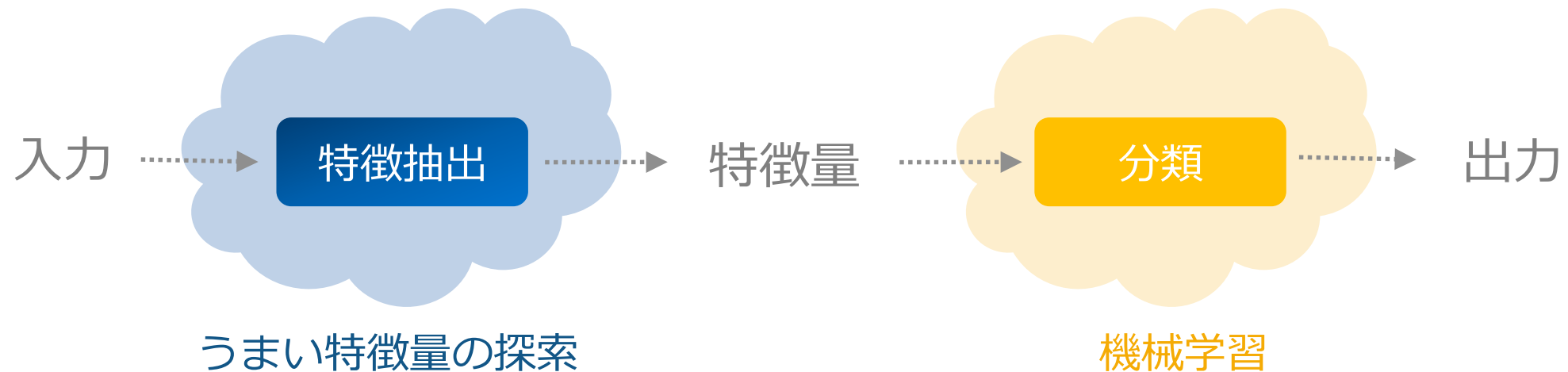
Accuracy : 0.84323

歩行	412	27	25	11	21	0
階段-上り	49	397	25	0	0	0
階段-下り	13	74	333	0	0	0
着席	1	17	0	382	86	5
直立	5	0	0	103	424	0
寝る	0	0	0	0	0	537
	歩行	階段-上り	階段-下り	着席	直立	寝る

混合行列（積層自己符号化器の場合）

特徴抽出と Deep Learning

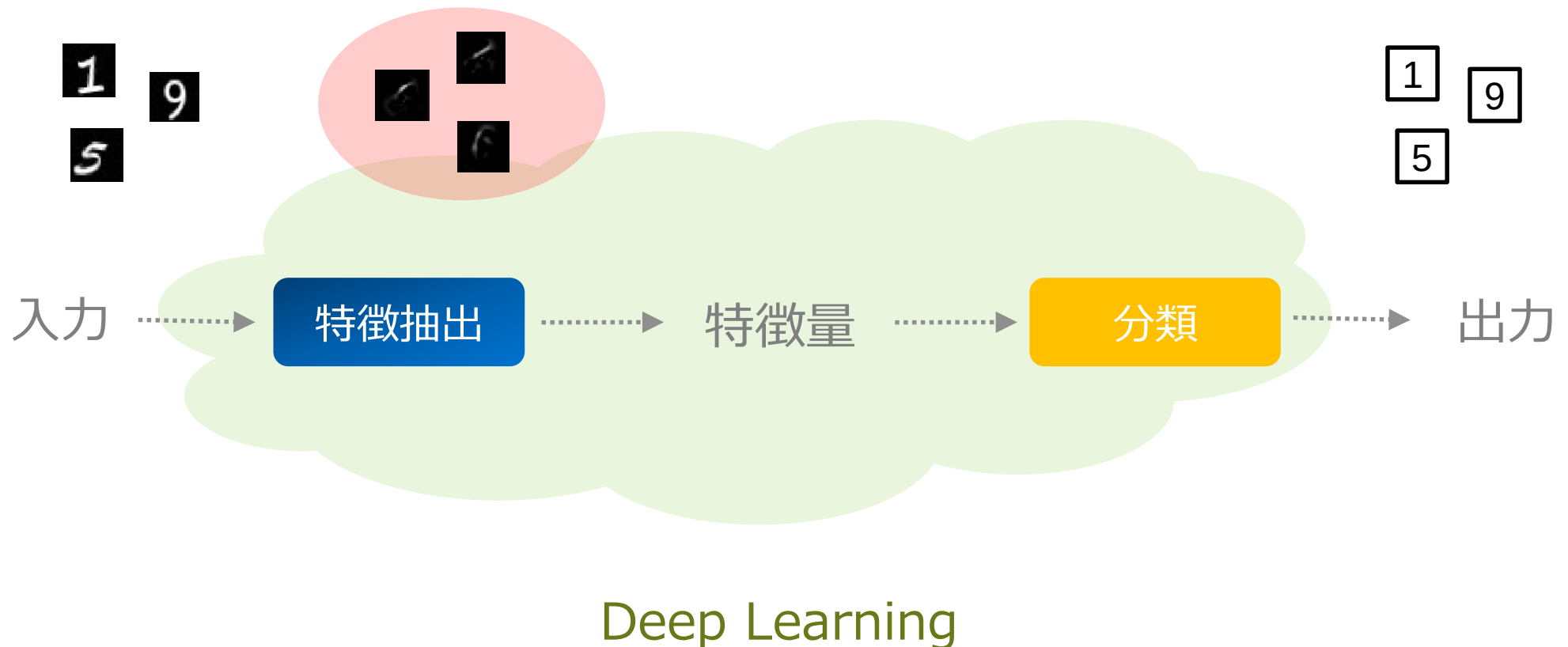
Deep Learning 以前は、特徴抽出の部分を試行錯誤で人間が探索していた



データに適した特徴量を見つけ出すことがよい性能を出すためのキーであった

特徴抽出と Deep Learning

Deep Learning では、特徴抽出の部分もデータから学習させることができる



LSTM

Long Short Time Memory

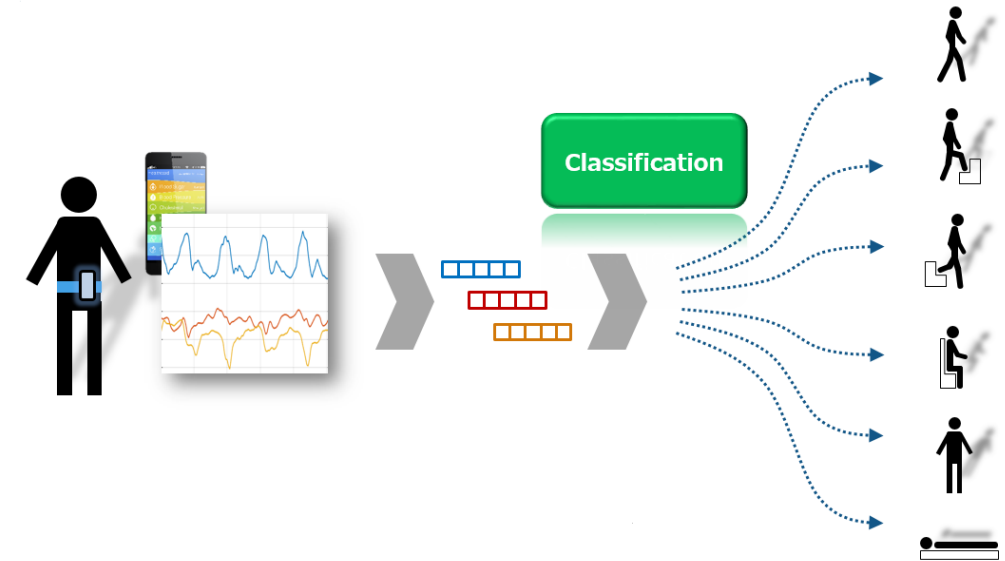
LSTM (Long Short Term Memory)

R2017b

3 軸の加速度センサーのデータで信号分類する際のネットワーク構築の例

```
layers = [ ...  
    sequenceInputLayer(3)  
    lstmLayer(50)  
    lstmLayer(50, 'OutputMode', 'last')  
    fullyConnectedLayer(6)  
    softmaxLayer  
    classificationLayer];
```

```
opts = trainingOptions('sgdm', 'MaxEpochs', 150);  
net = trainNetwork(XTrain, TTrain, layers, opts);
```



入力層・LSTM層・全結合層などの
層を積み上げて定義

学習率や最大反復数などを定義して
学習の関数を呼び出す

【例題】人の活動状態の識別

LSTM (Long Short Term Memory) にすることで
更に分類性能が 5% 程度改善！

Accuracy : 0.84323

歩行	412	27	25	11	21	0
階段-上り	49	397	25	0	0	0
階段-下り	13	74	333	0	0	0
着席	1	17	0	382	86	5
直立	5	0	0	103	424	0
寝る	0	0	0	0	0	537
	歩行	階段-上り	階段-下り	着席	直立	寝る

混合行列（積層自己符号化器の場合）

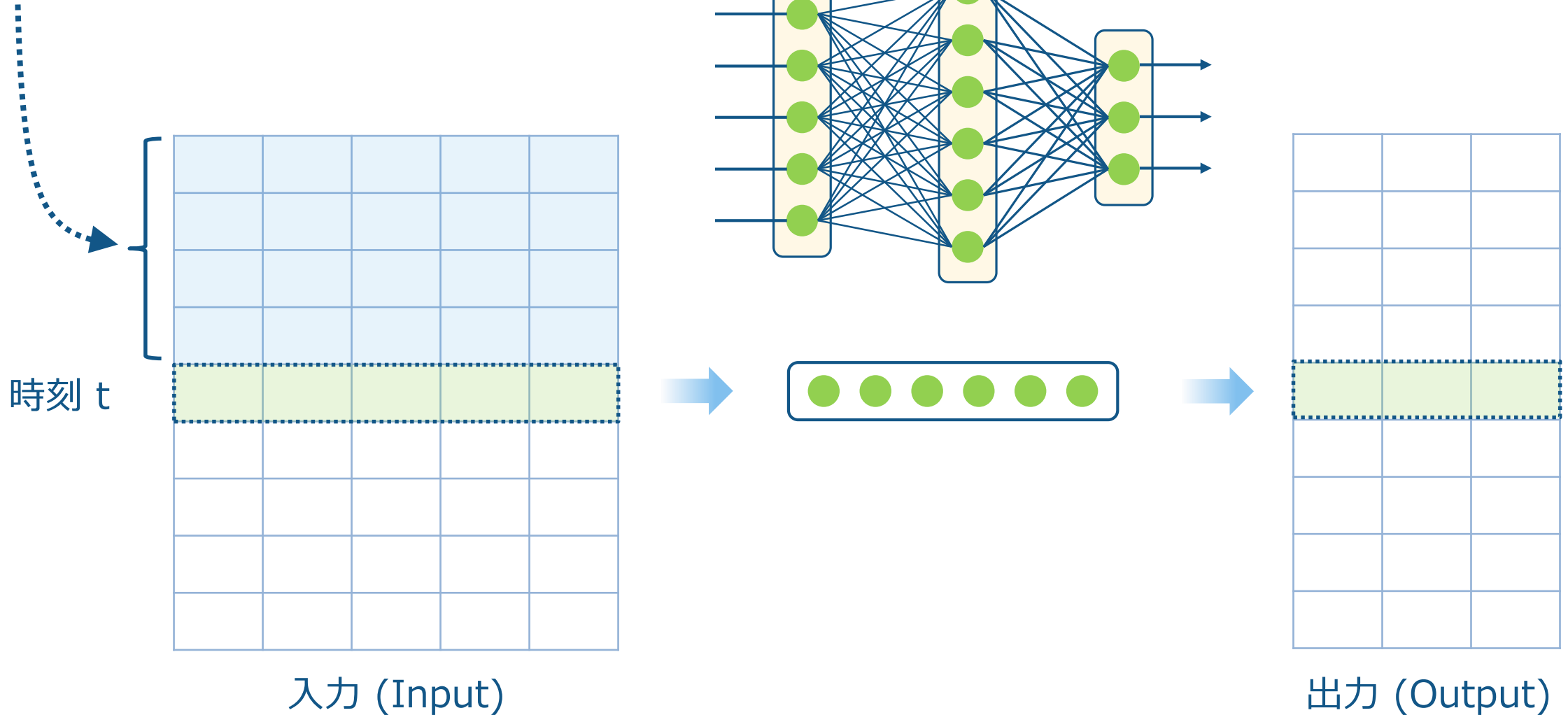
Accuracy : 0.89752

歩行	488	8	0	0	0	0
階段-上り	35	426	10	0	0	0
階段-下り	3	30	387	0	0	0
着席	0	0	0	373	114	4
直立	1	0	0	97	434	0
寝る	0	0	0	0	0	537
	歩行	階段-上り	階段-下り	着席	直立	寝る

混合行列（LSTM の場合）

Recurrent Neural Network とは？

過去のデータを現時点の予測にどう活かすか？



Recurrent Neural Network とは？

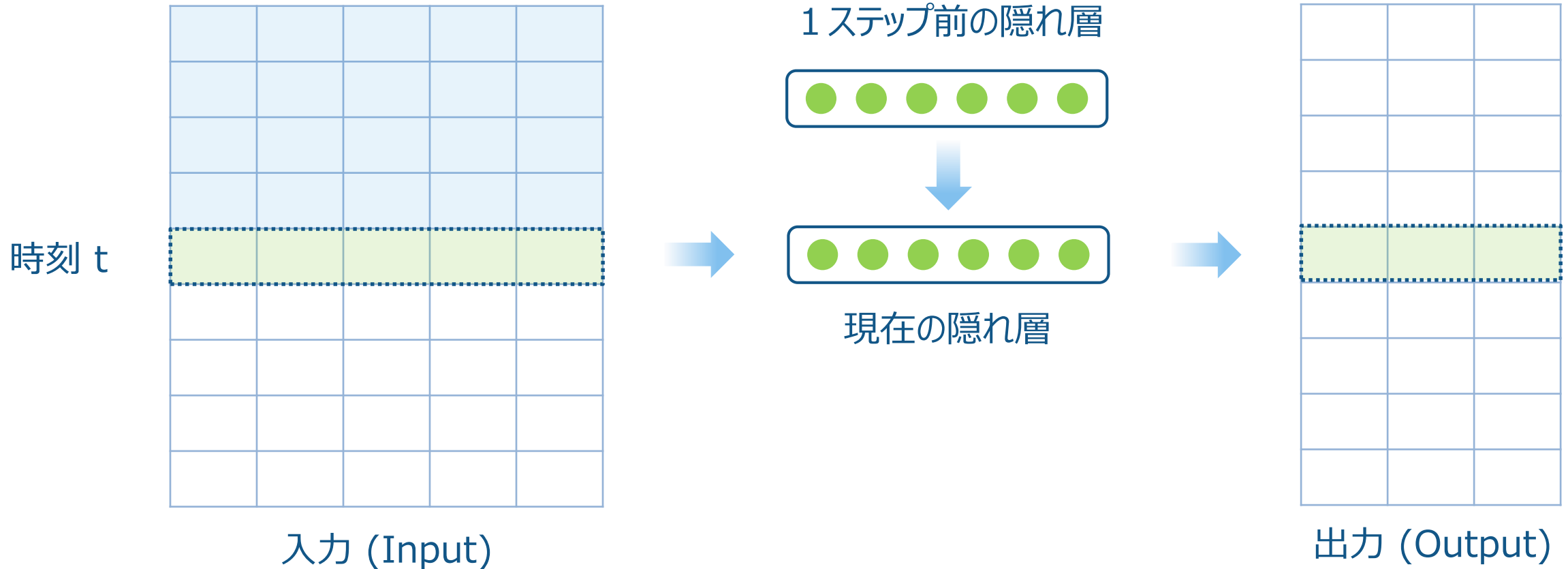
現時刻の入力の情報

+

1 ステップ前の隠れ層の情報



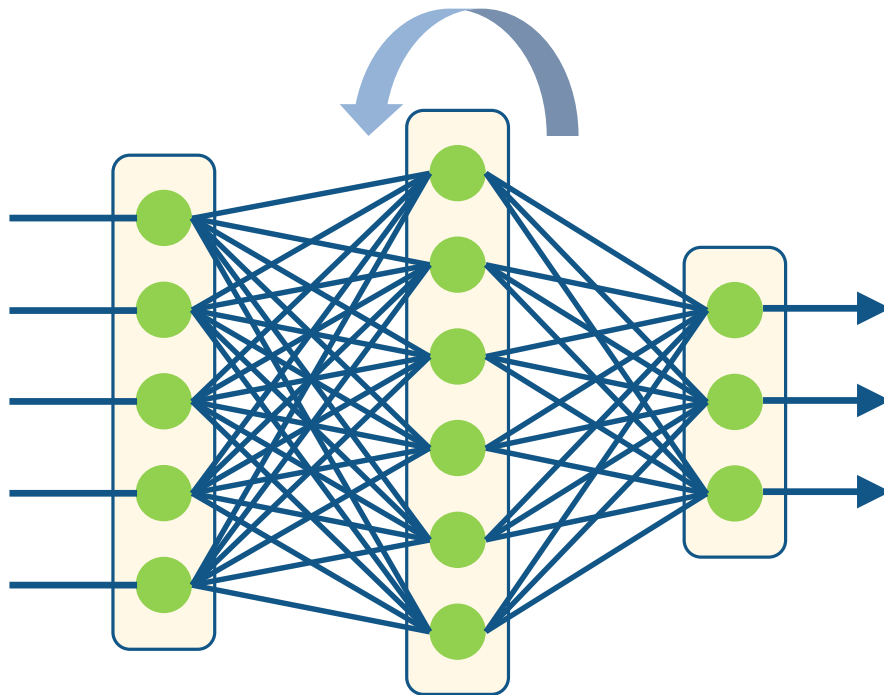
過去情報も含む予測



Recurrent Neural Network とは？

R2017b

1 ステップ前の隠れ層の状態をフィードバックさせることで過去の情報を予測に活かせるようにしたネットワーク



エルマン型 RNN

長期間の依存関係を
モデル化するのが不得意

進化！

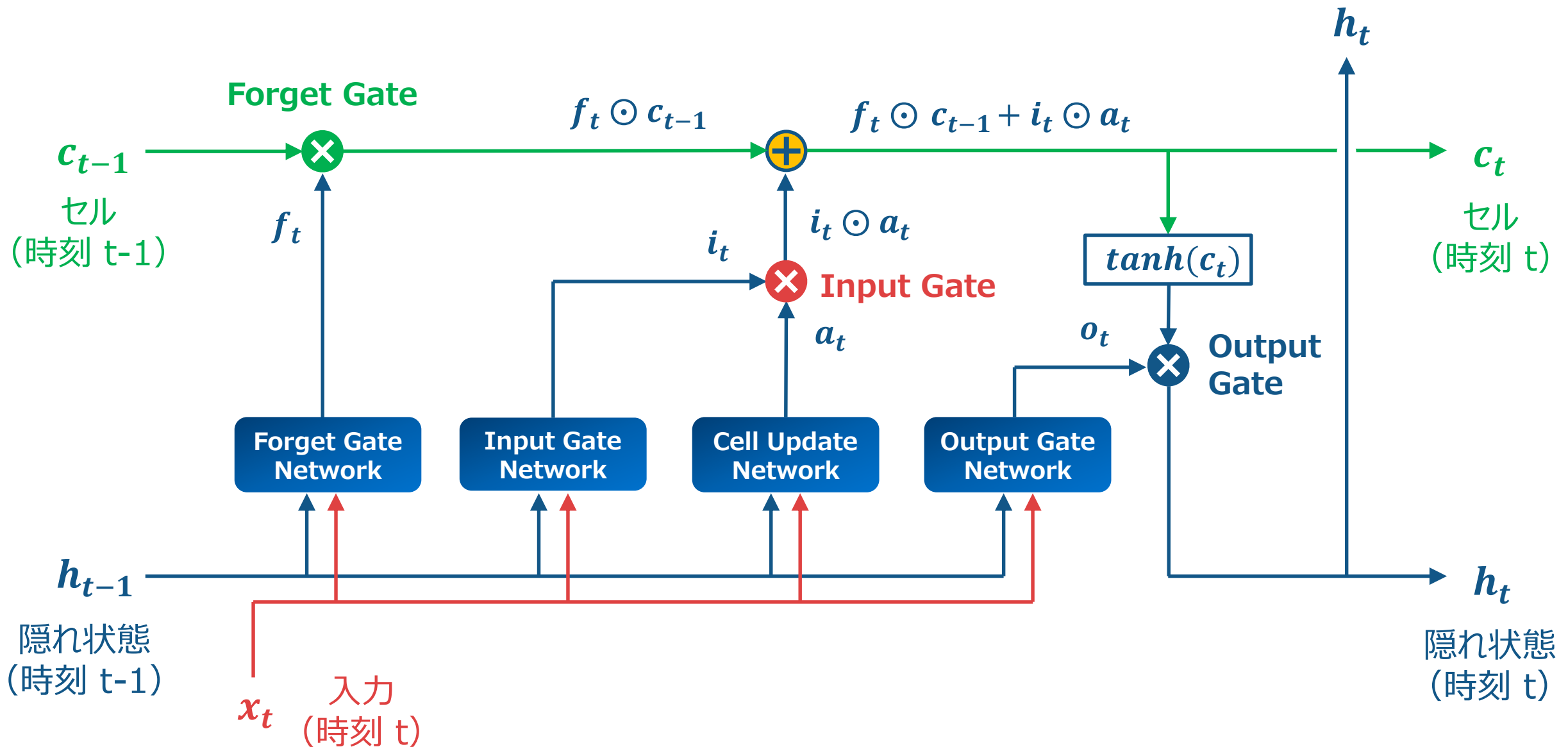


LSTM

Long Short Time Memory

- 進化版RNN（ゲート付きRNN）
- メモリセルにより長期の情報を保持できる
- 時間的に離れた依存関係のモデル化も得意

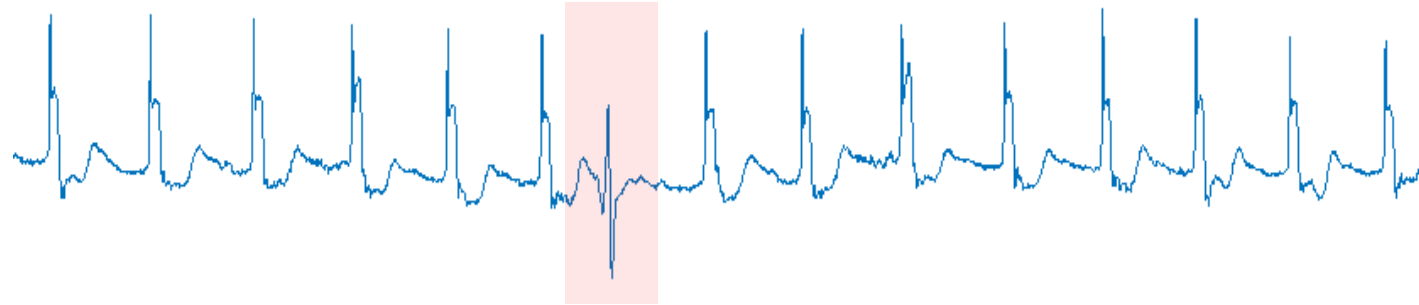
LSTM (Long Short Term Memory) とは？



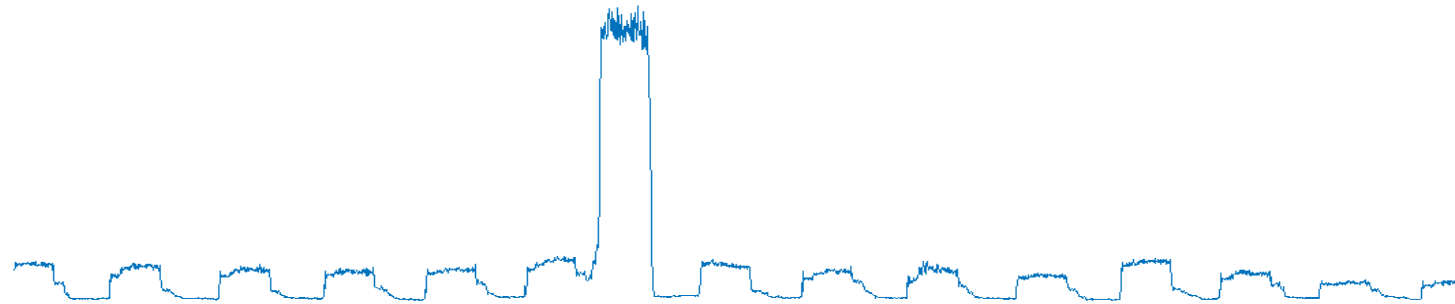
時系列データの異常検知

センサーデータからの異常検出

心電図



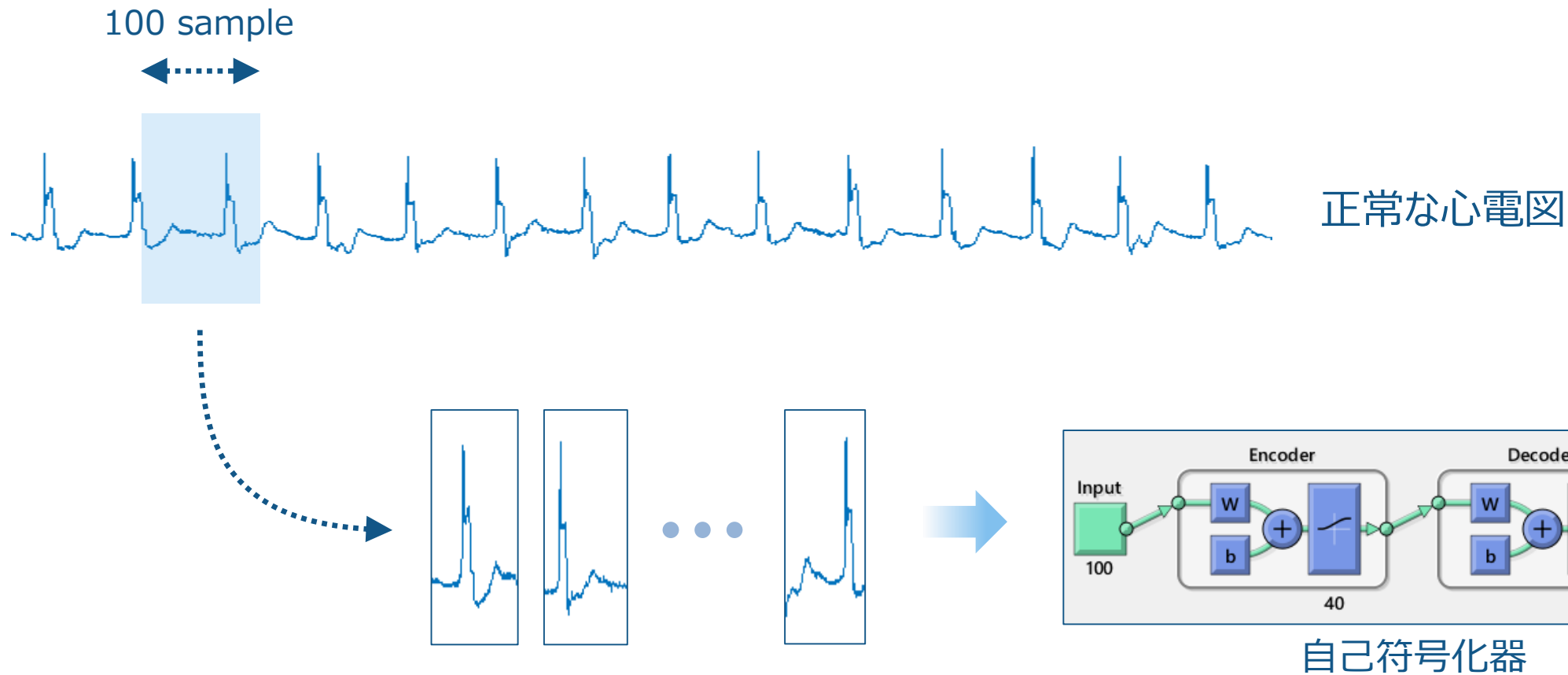
異常度



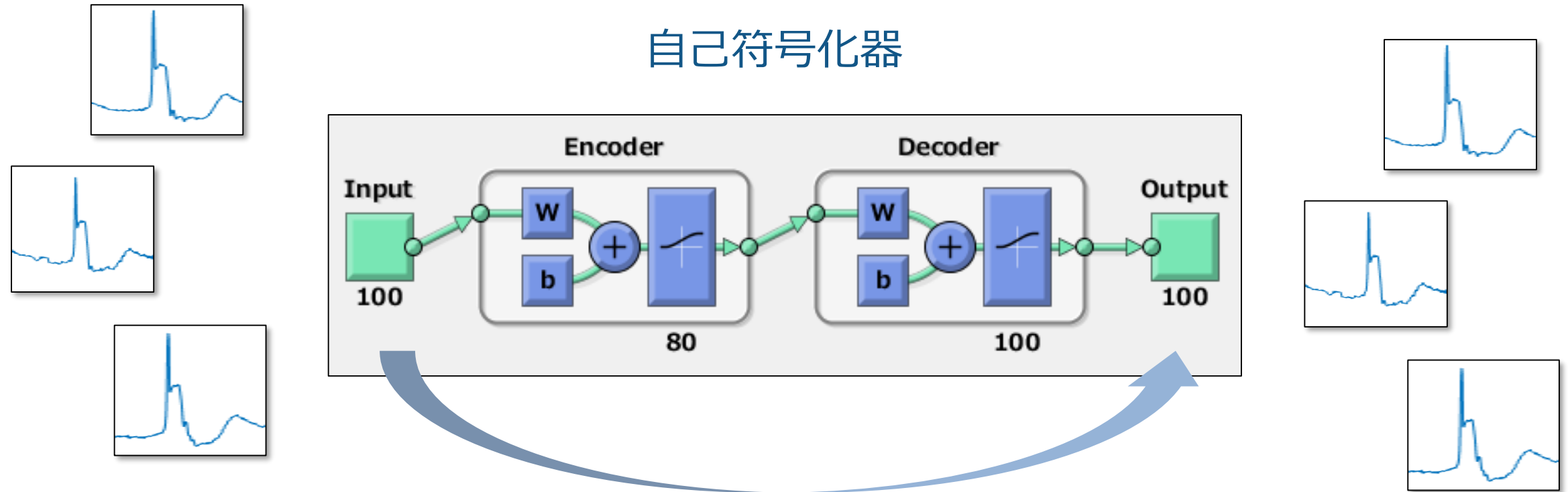
Keogh, E., Lin, J. and Fu, A. : HOT SAX : Efficiently Finding the Most Unusual Time Series Subsequence, in Proceedings of the Fifth IEEE International Conference on Data Mining, ICDM 05, pp.226-233

自己符号化器による特徴抽出

正常な心電図から部分時系列を生成して、自己符号化器に学習させる



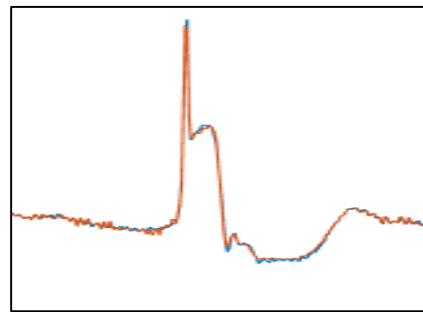
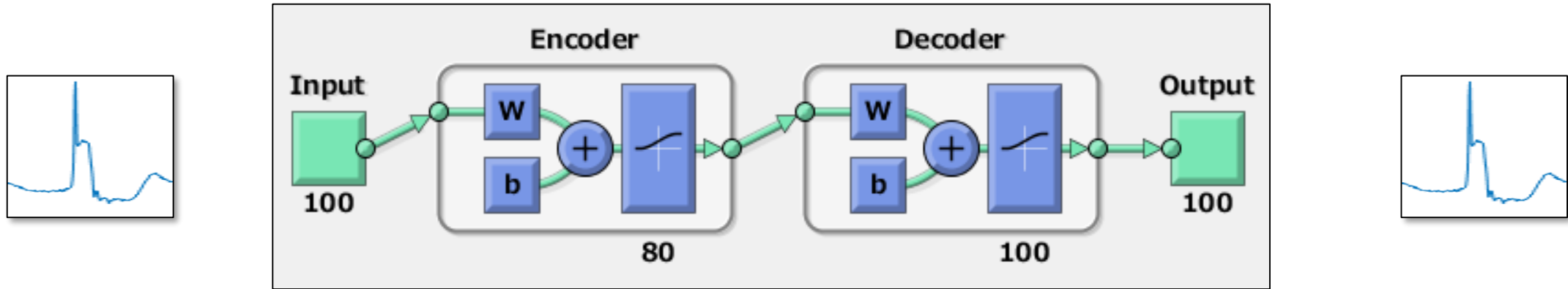
自己符号化器による異常検出



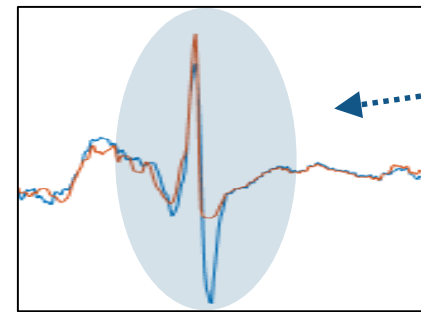
元の正常な波形を再生するようにウェイトを学習する

自己符号化器による異常検出

正常な波形の場合は、きれいに元の波形を再生するが、異常な波形の場合はうまく再生できない！



正常な波形の場合



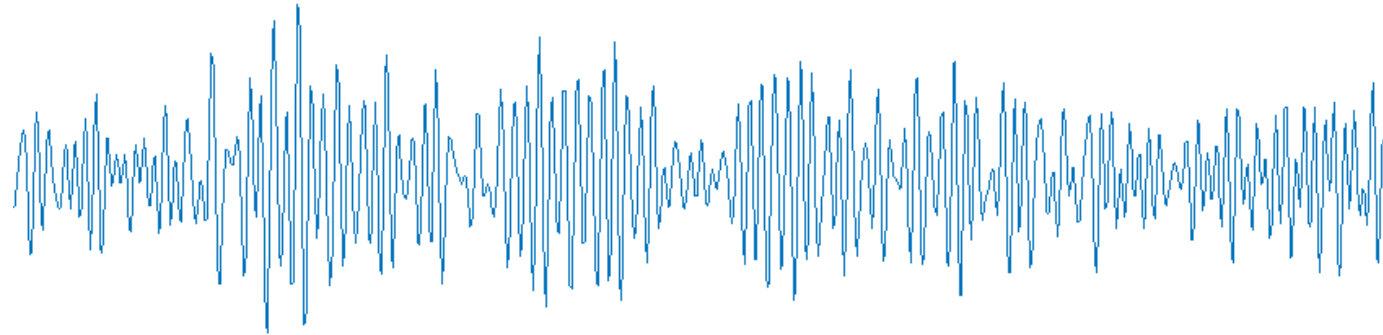
異常な波形の場合

異常な波形は正常な
波形に戻そうとする

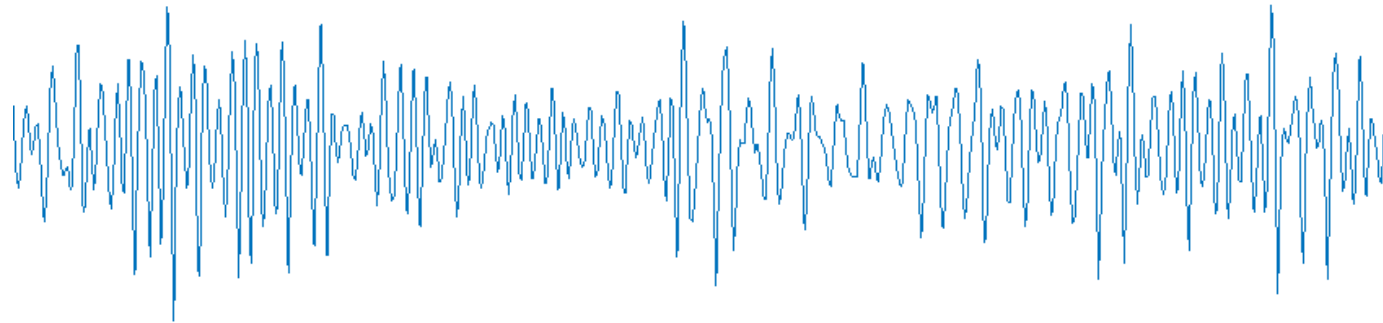
構造物の異常検出

ある構造物に取り付けられた振動データの比較

正常



異常



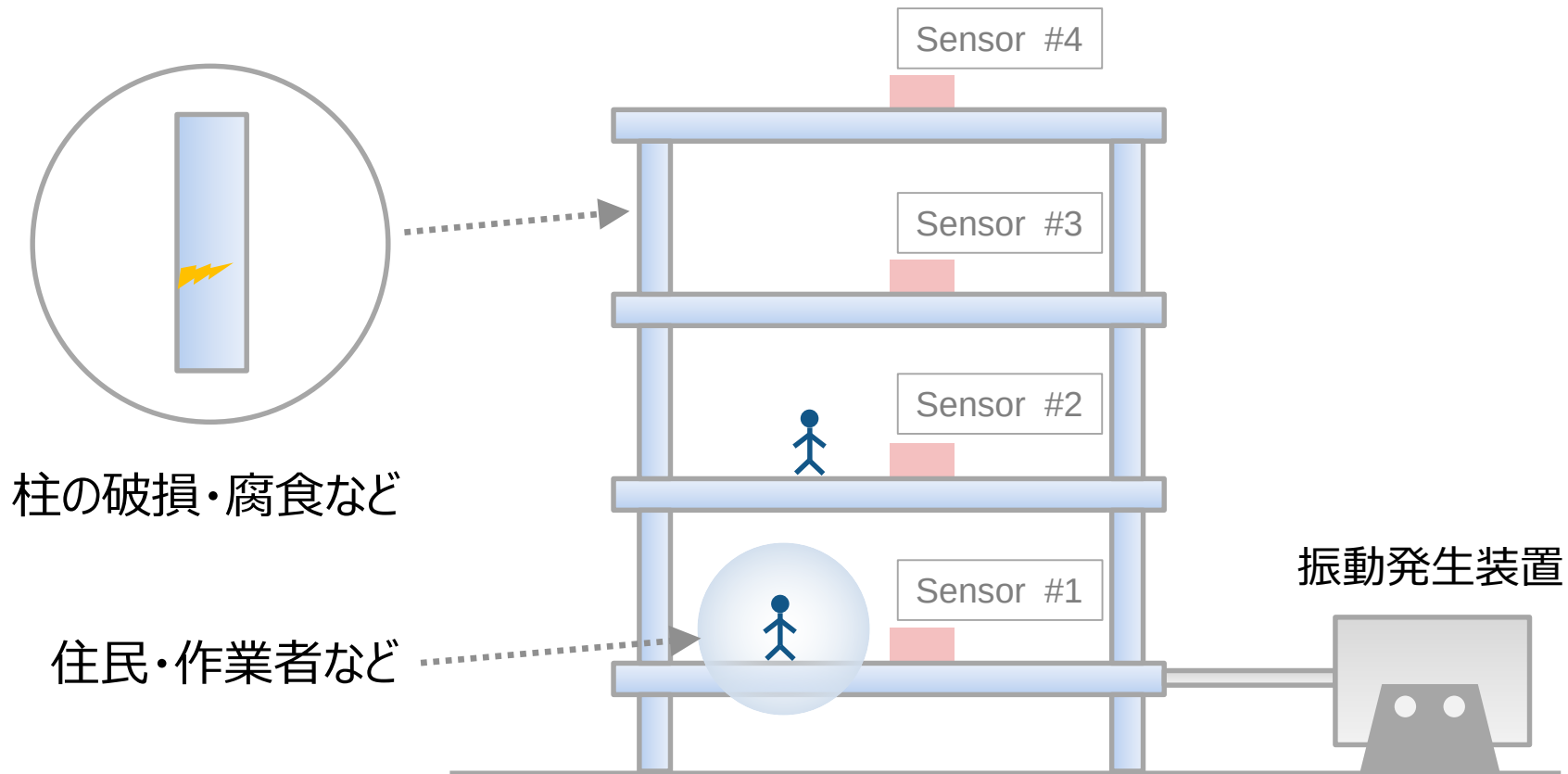
【例題】 構造ヘルスマニタリング



- 実験装置
 - 3 階建ての建物の模型
 - 4 ヶ所の加速度センサー
 - 1 階部分に振動発生装置
 - 人工的な損傷状態の発生
- 実験目的
 - センサーによる異常検出

Reference: Figueiredo, E., Park, G., Figueiras, J., Farrar, C., & Worden, K. (2009). Structural Health Monitoring Algorithm Comparisons using Standard Data Sets. Los Alamos National Laboratory Report: LA-14393

【例題】 構造ヘルスマニタリング



Reference: Figueiredo, E., Park, G., Figueiras, J., Farrar, C., & Worden, K. (2009). Structural Health Monitoring Algorithm Comparisons using Standard Data Sets. Los Alamos National Laboratory Report: LA-14393

【例題】 構造物ヘルスマモニタリング

構造物の模型において、次の条件を変化させて幾つかの条件を作り出した。

- 各階への重量（Mass）
- 柱の硬さの低減（Stiffness reduction）

State	Description
#1	Baseline condition
#2	Mass = 1.2 kg at the base
#3	Mass = 1.2 kg on the 1st floor
#4	87.5% stiffness reduction in column 1BD
#5	87.5% stiffness reduction in column 1AD and 1BD
#6	87.5% stiffness reduction in column 2BD
#7	87.5% stiffness reduction in column 2AD and 2BD
#8	87.5% stiffness reduction in column 3BD
#9	87.5% stiffness reduction in column 3AD and 3BD

上記 9 パターンのそれぞれで 10 回の実験を繰り返し、合計 90 回の実験を行ったものとします

【例題】 構造物ヘルスマニタリング

State #1 を正常データとして学習させて、異常データが検知できるか調べたい

State	Description	
#1	Baseline condition	}→ 正常データ
#2	Mass = 1.2 kg at the base	
#3	Mass = 1.2 kg on the 1st floor	
#4	87.5% stiffness reduction in column 1BD	}→ 異常データ
#5	87.5% stiffness reduction in column 1AD and 1BD	
#6	87.5% stiffness reduction in column 2BD	
#7	87.5% stiffness reduction in column 2AD and 2BD	
#8	87.5% stiffness reduction in column 3BD	
#9	87.5% stiffness reduction in column 3AD and 3BD	

正常データを学習させる → 異常データを検出できるか？

SHMTools

Los Alamos National Laboratory

Los Alamos National Laboratory National Security Education Center search site

Engineering Institute

Addressing national needs by fostering specialized recruiting and strategic partnerships

EDUCATIONAL PROGRAMS SHM DATA SETS AND SOFTWARE RESEARCH PROJECTS PUBLICATIONS

NSEC » Engineering Institute » SHM Data Sets and Software

SHM Data Sets and Software

SHMTools is a MATLAB package that facilitates the construction of structural health monitoring (SHM) processes.

CONTACT
Institute Director
Charles Farrar
(505) 665-0860
[Email](#)

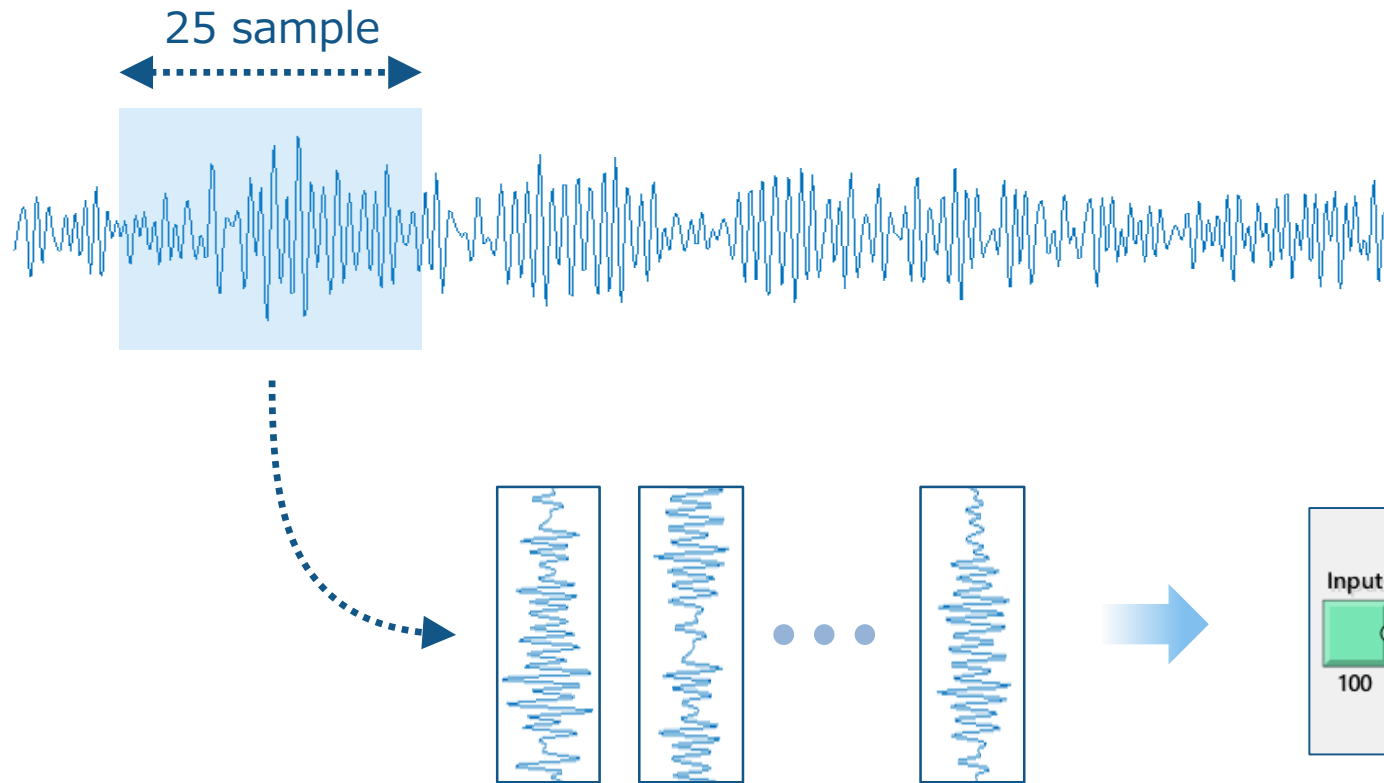
SHMTools Software (0.3.0 Beta)
The package provides a set of functions organized into modules according to the three primary stages of Structural Health Monitoring: Data Acquisition, Feature Extraction, and Feature Classification. A modular function design and set of standardized parameter formats make it easy to assemble and test customized SHM processes.

Related Links
[Getting Started with SHMTools \(pdf\)](#)
[mFUSE Help Manual \(pdf\)](#)

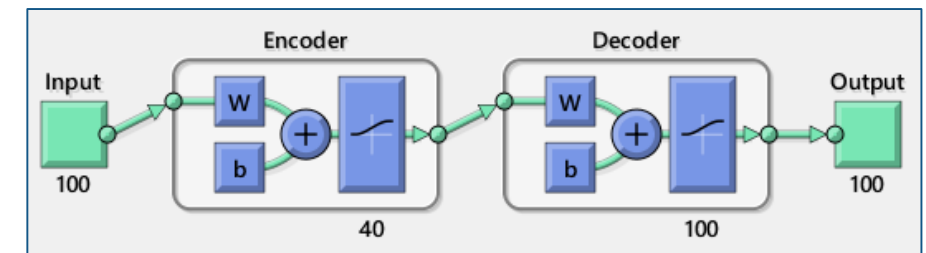
<http://www.lanl.gov/projects/national-security-education-center/engineering/software/shm-data-sets-and-software.php>

自己符号化器による異常検出

正常な振動データから部分時系列を生成して、自己符号化器に学習させる



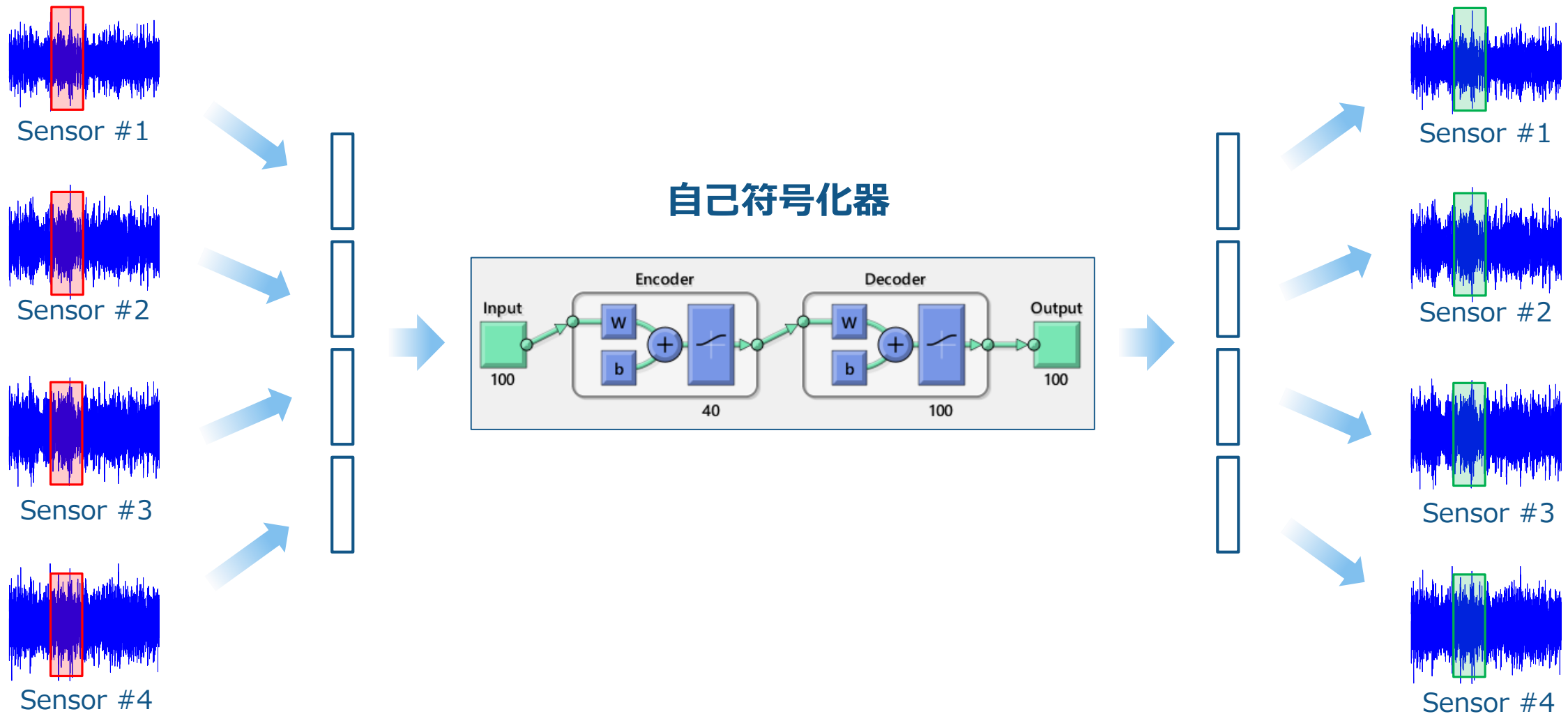
正常な振動データ



自己符号化器

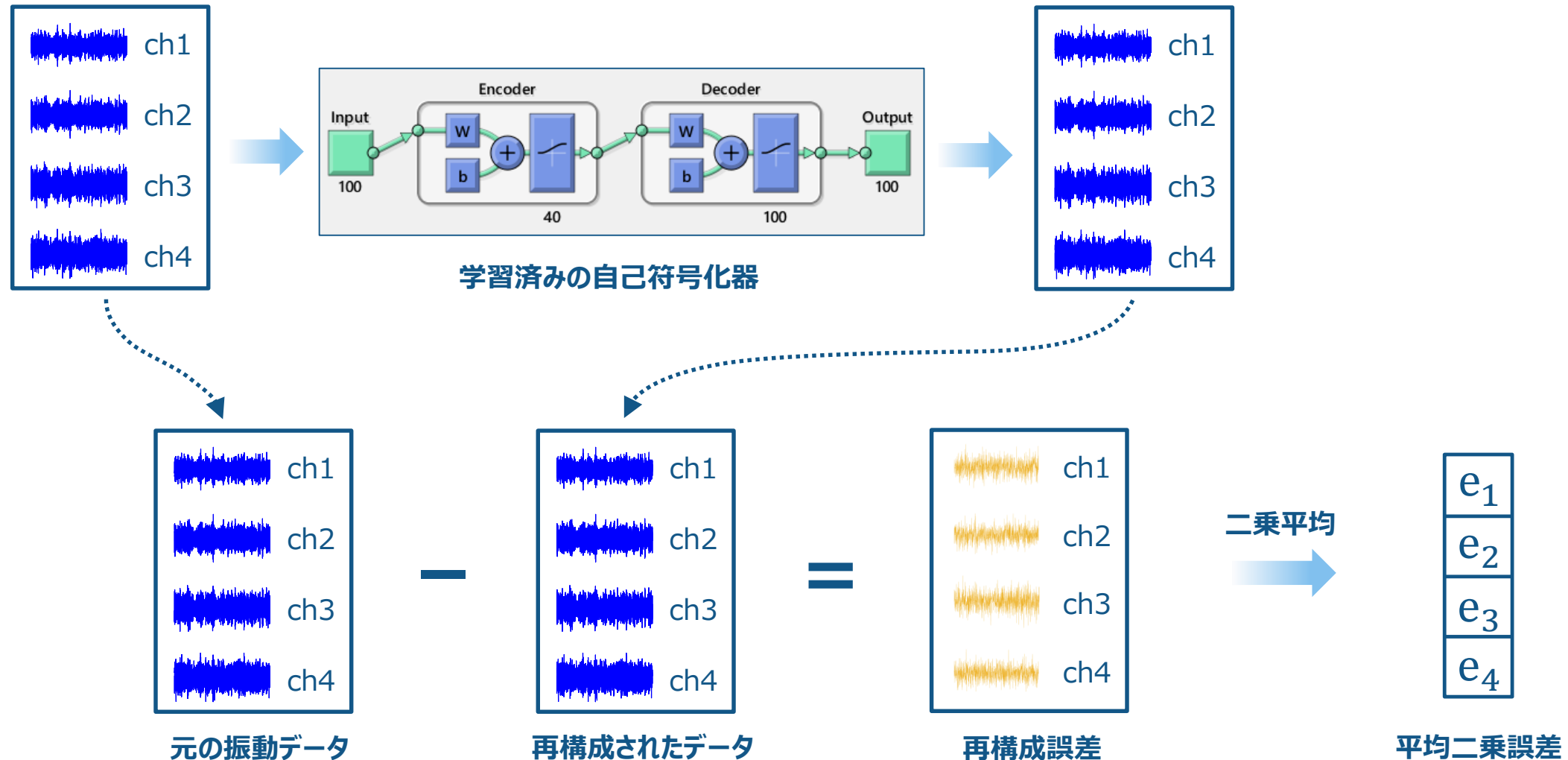
自己符号化器による異常検出

正常な場合のデータを自己符号化器に学習させる

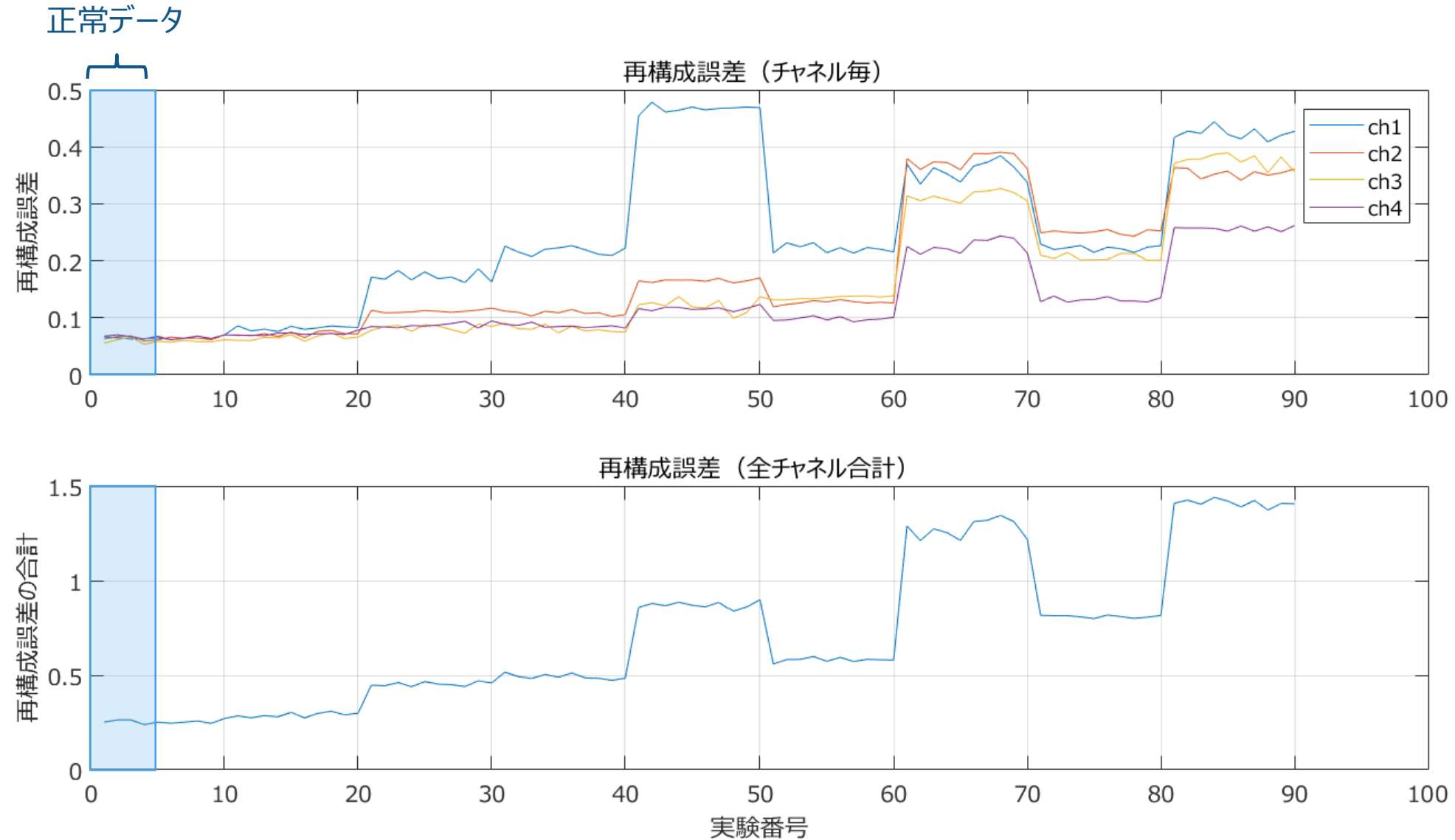


自己符号化器による異常検出

異常を含む可能性のあるデータを自己符号化器に通して、再構成誤差を算出する



自己符号化器による異常検出

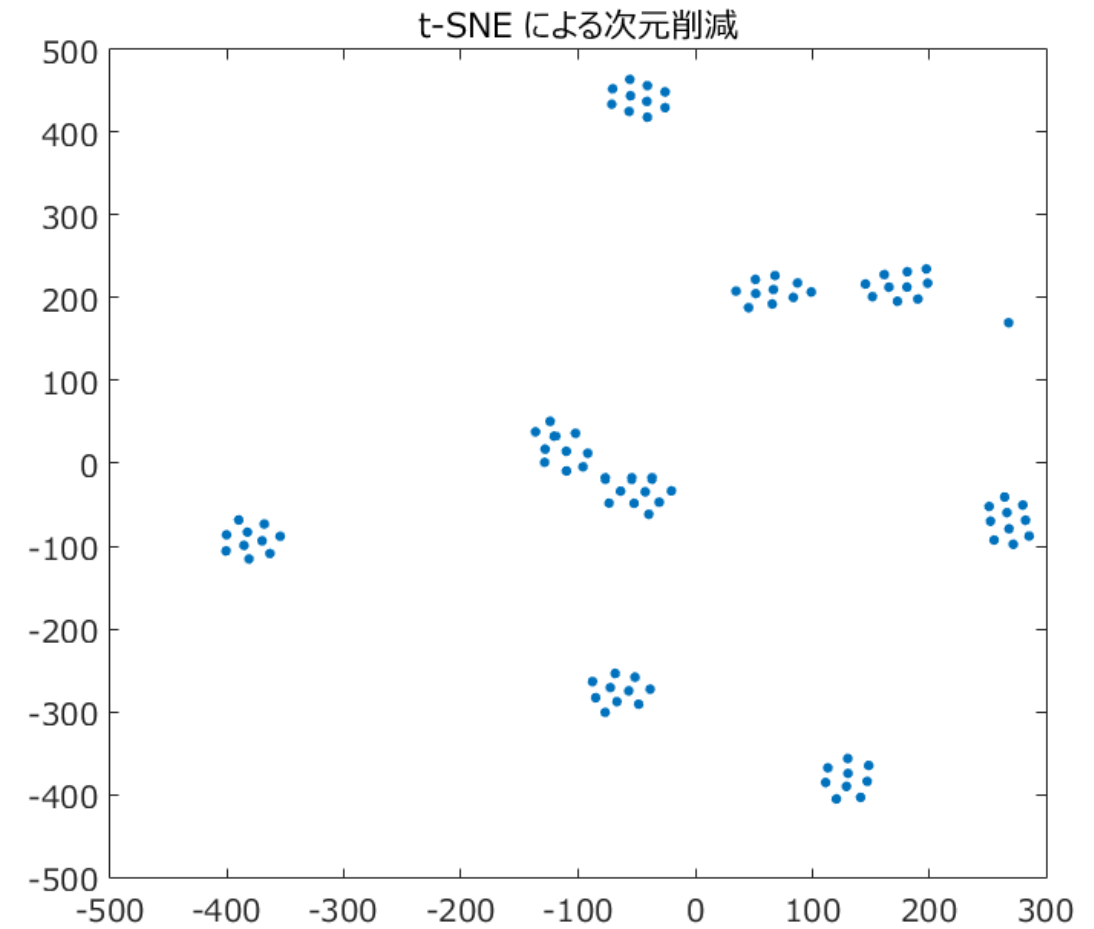
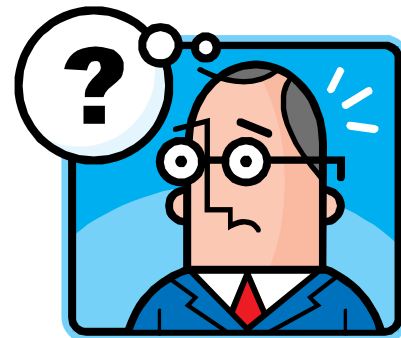


t-SNE による異常データのクラスタリング

検出された異常の種類を識別することはできるだろうか？

State	Description
#1	Baseline condition
#2	Mass = 1.2 kg at the base
#3	Mass = 1.2 kg on the 1st floor
#4	87.5% stiffness reduction in column 1BD
#5	87.5% stiffness reduction in column 1AD and 1BD
#6	87.5% stiffness reduction in column 2BD
#7	87.5% stiffness reduction in column 2AD and 2BD
#8	87.5% stiffness reduction in column 3BD
#9	87.5% stiffness reduction in column 3AD and 3BD

どのような異常の種類があるだろうか？

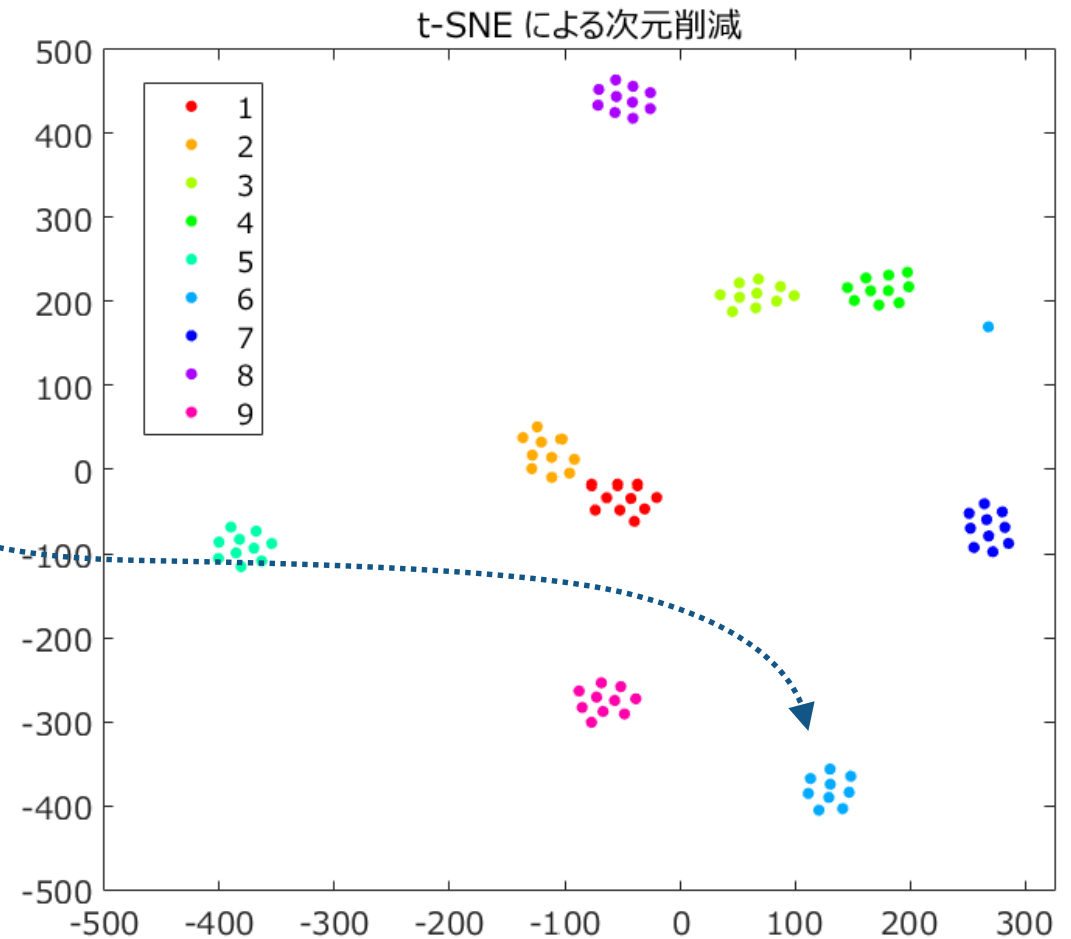


t-SNE による異常データのクラスタリング

実際の異常の種類で色分けすると、次のような可視化を行うことができる

State	Description
#1	Baseline condition
#2	Mass = 1.2 kg at the base
#3	Mass = 1.2 kg on the 1st floor
#4	87.5% stiffness reduction in column 1BD
#5	87.5% stiffness reduction in column 1AD and 1BD
#6	87.5% stiffness reduction in column 2BD
#7	87.5% stiffness reduction in column 2AD and 2BD
#8	87.5% stiffness reduction in column 3BD
#9	87.5% stiffness reduction in column 3AD and 3BD

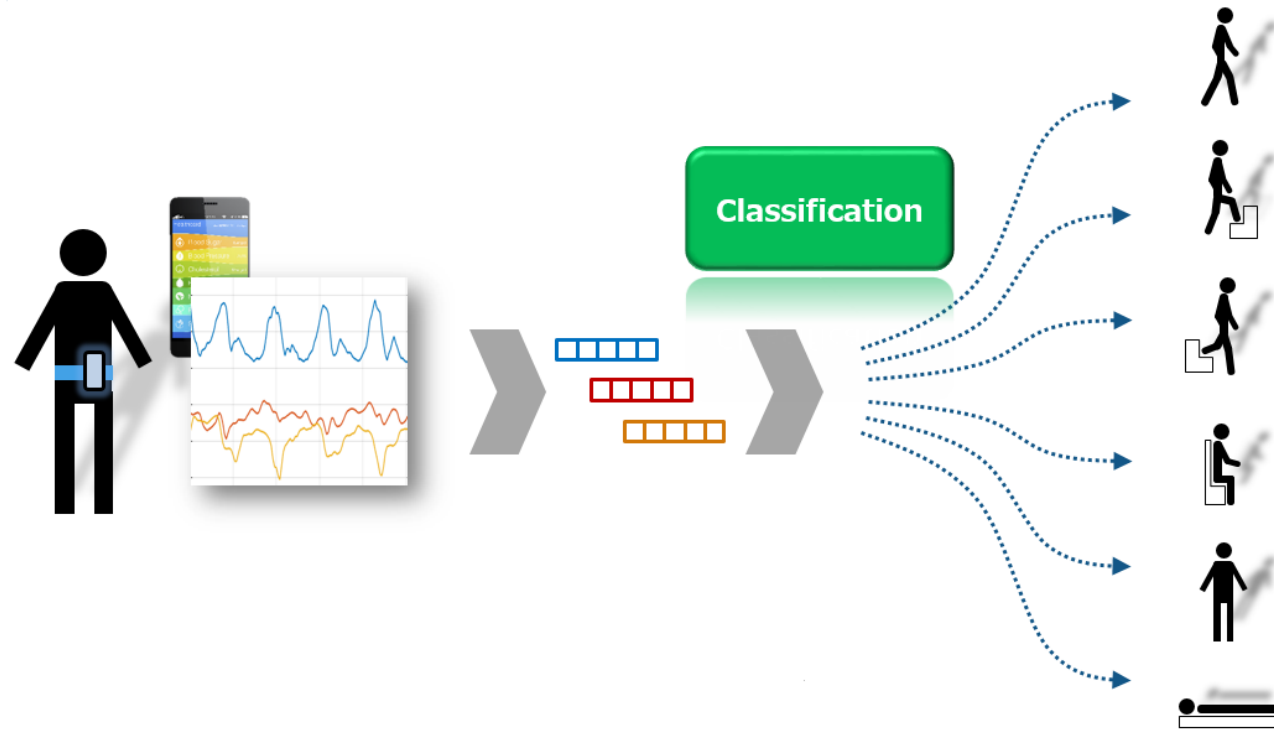
それぞれの異常がひとつの
島に対応しているらしいな！



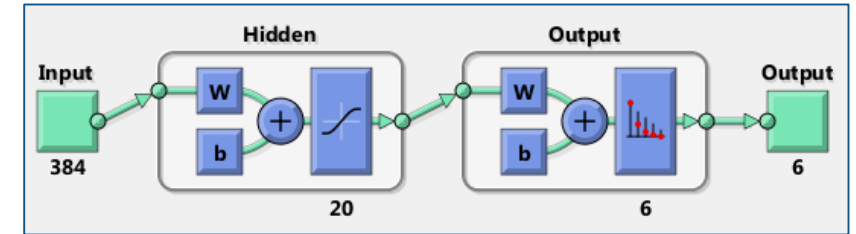
デモのまとめ

【例題】人の活動状態の識別

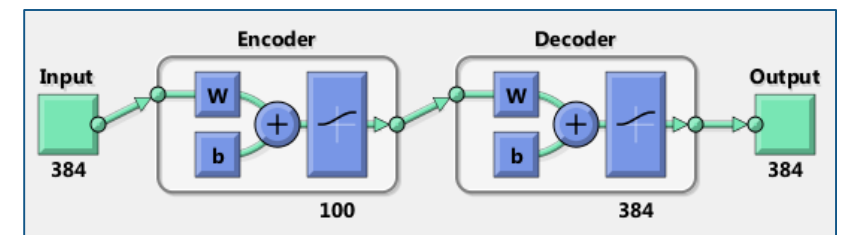
ニューラルネットを分類と次元削減の両方に利用



部分時系列からシンプルに分類



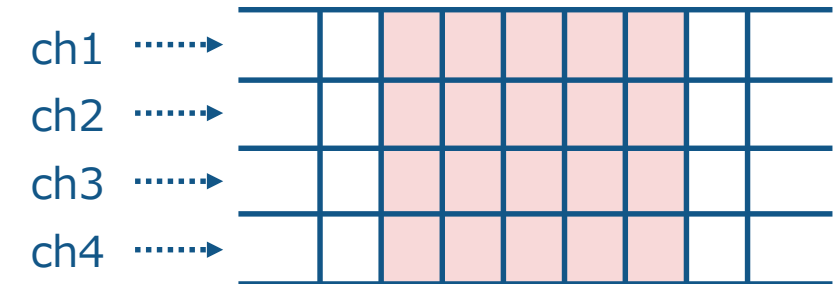
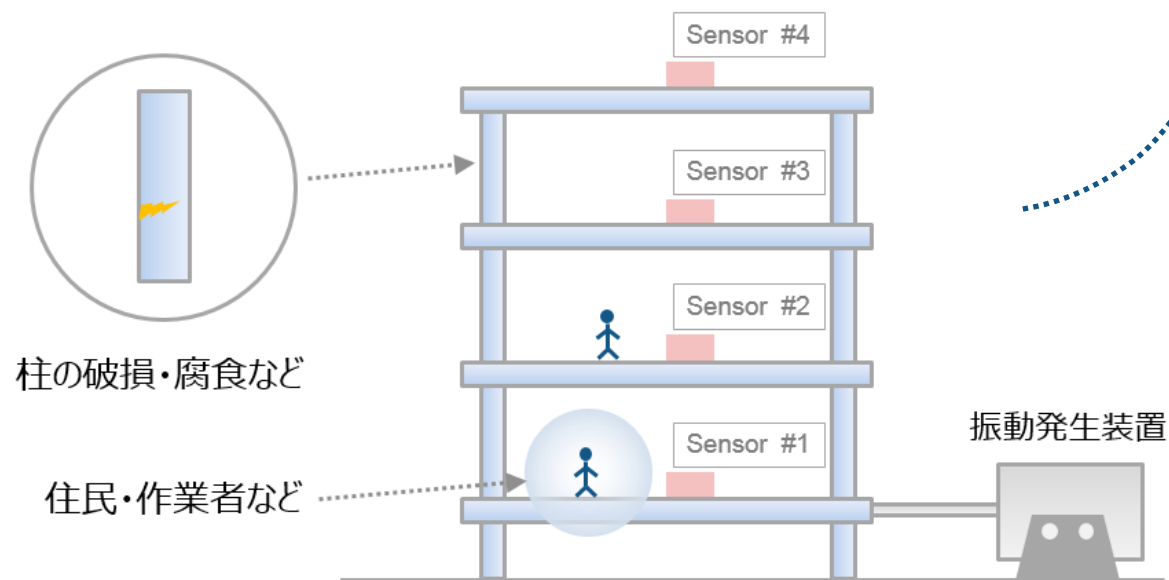
エンコーダ部分を次元削減に利用



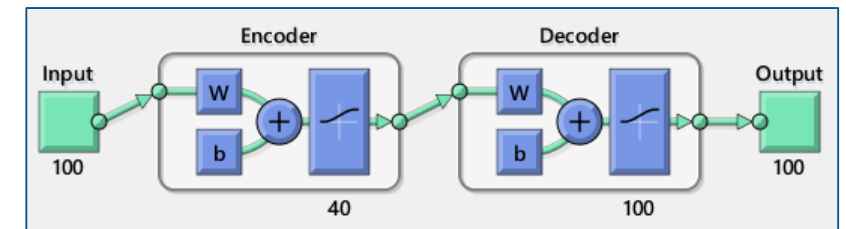
➡ 元々のデータで分類がうまくいくようなケースでは次元削減で性能向上が狙える可能性が！

【例題】構造ヘルスマニタリング

正常なセンサーデータが持つ「データの関係性」をモデル化



部分時系列として切り出される
データの関係性をモデル化する



➡ 「データの関係性」の崩れを「再構成誤差」として検出することが可能となる！

システムへの展開・Cコード生成

センサーデータ解析システムにおける課題

■ 端末の問題

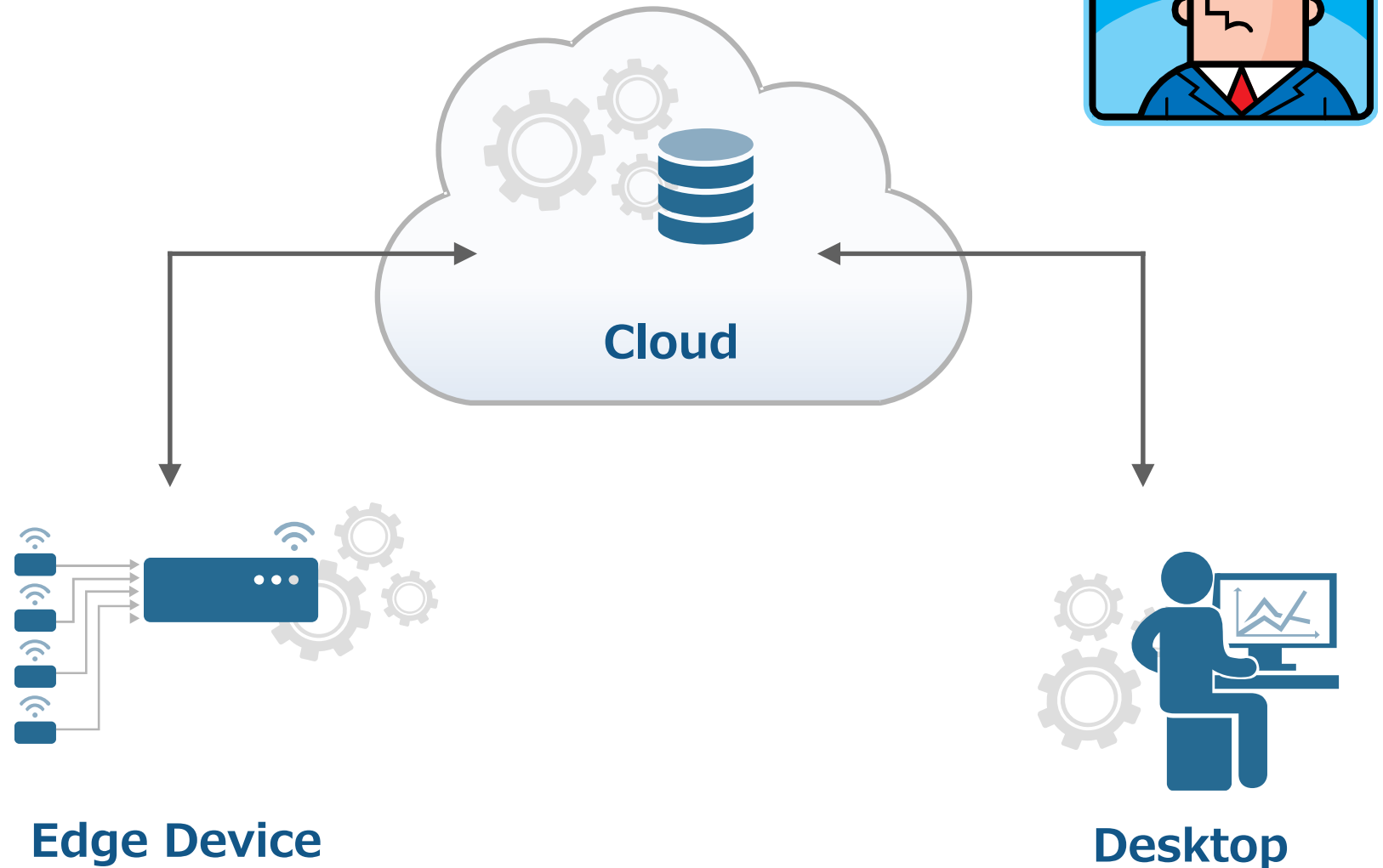
- ・ データを全部送るか？
- ・ 特徴量のみを送るか？
- ・ 異常なデータのみ送るか？
- ・ 通信モジュールの消費電力は？
- ・ 演算速度と消費電力は？

■ インフラの問題

- ・ 必要な帯域は確保されているか？
- ・ ネットワークは安定しているか？
- ・ 電力供給は安定しているか？
- ・ 通信量に対するコストは？

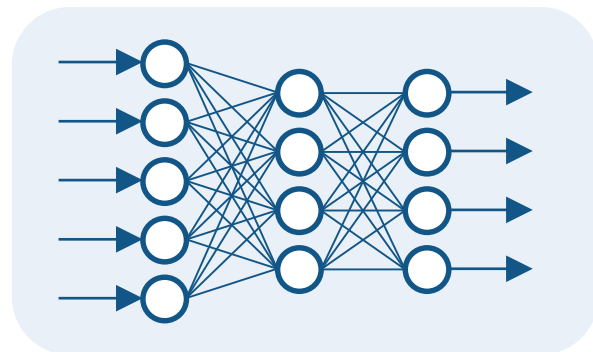
■ 開発側の問題

- ・ アルゴリズムの検証し易さ
- ・ 組み込みコードの開発コスト
- ・ サーバ側コードの開発コスト



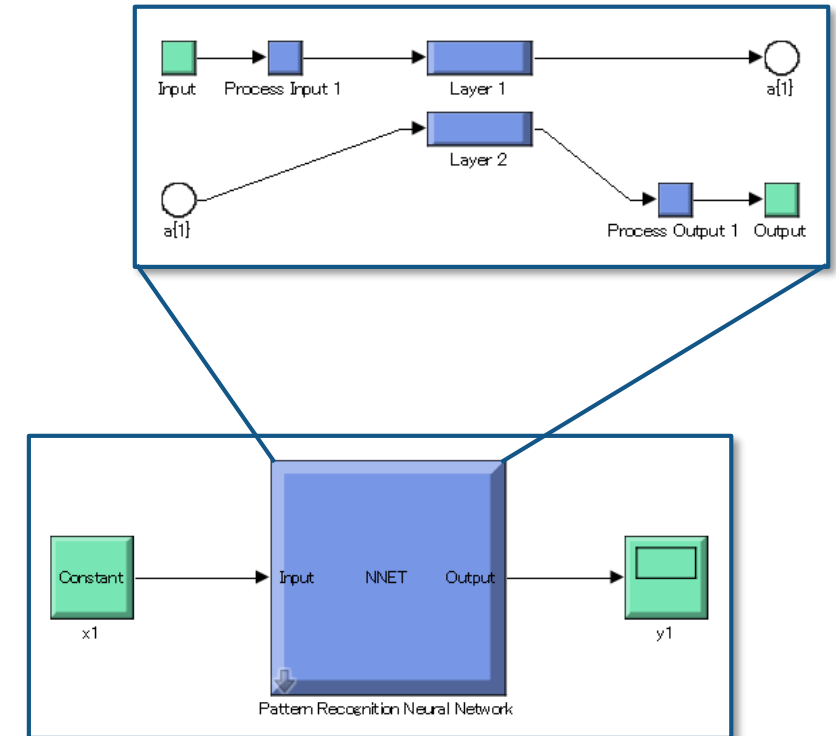
学習済み自己符号化器の活用

- Simulink ブロック への変換
- MATLAB 関数への変換



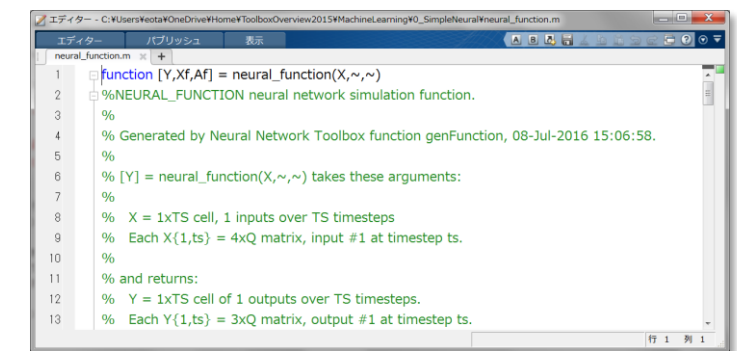
学習済み自己符号化器

Simulinkブロック生成
(関数 : generateSimulink)



Simulinkブロック

MATLABコード生成
(関数 : generateFunction)

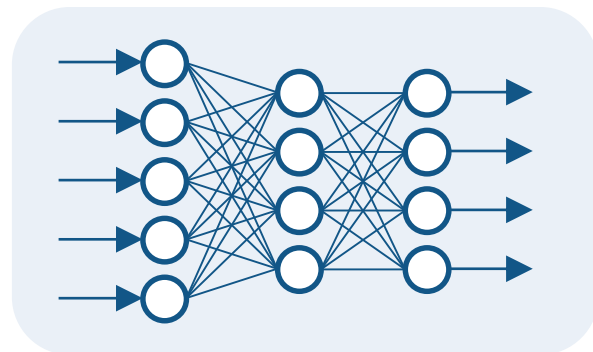


```
1 function [Y,Xf,Af] = neural_function(X,~,~)
2 %NEURAL_FUNCTION neural network simulation function.
3 %
4 % Generated by Neural Network Toolbox function genFunction, 08-Jul-2016 15:06:58.
5 %
6 % [Y] = neural_function(X,~,~) takes these arguments:
7 %
8 % X = 1xTS cell, 1 inputs over TS timesteps
9 % Each X{1,ts} = 4xQ matrix, input #1 at timestep ts.
10 %
11 % and returns:
12 % Y = 1xTS cell of 1 outputs over TS timesteps.
13 % Each Y{1,ts} = 3xQ matrix, output #1 at timestep ts.
```

MATLAB関数

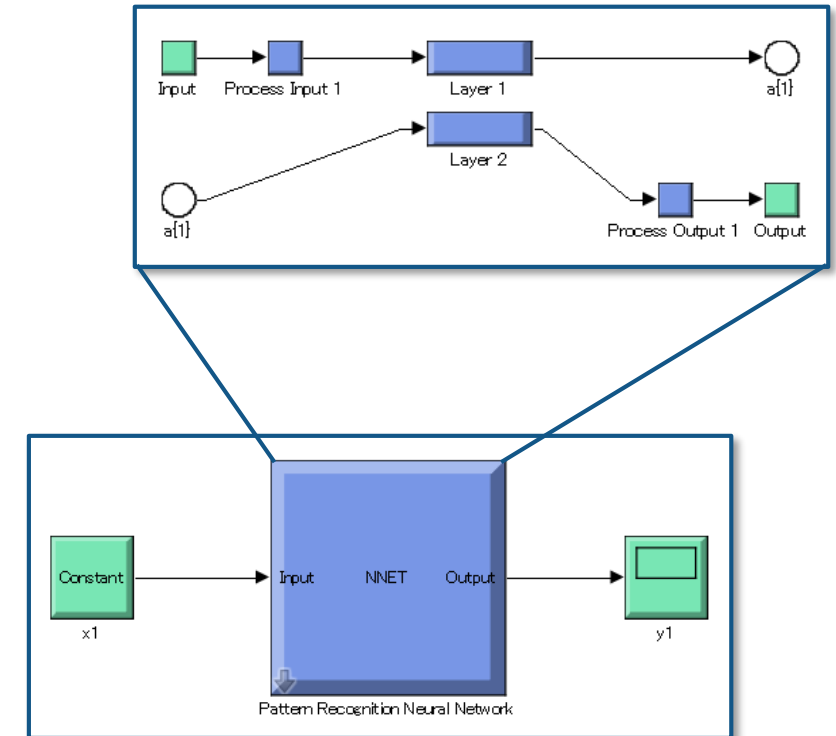
学習済みニューラルネットの活用

- Simulink ブロック への変換
- MATLAB 関数への変換



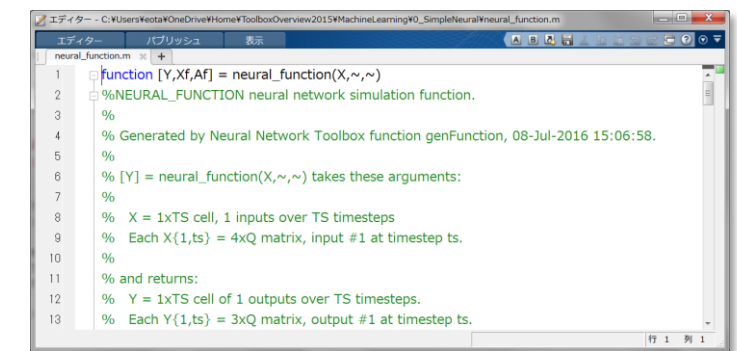
学習済みニューラルネット

Simulinkブロック生成
(関数 : gensim)



Simulinkブロック

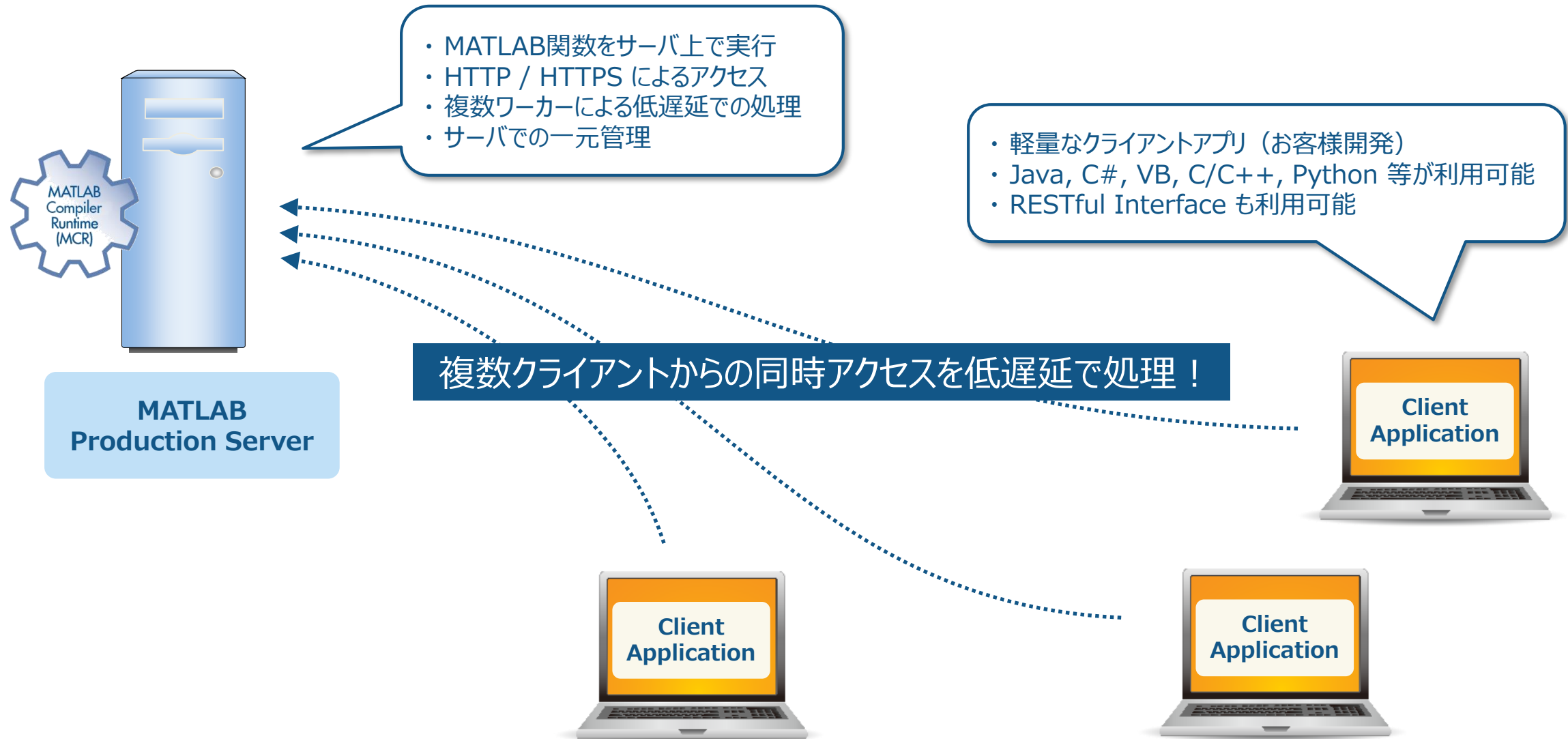
MATLABコード生成
(関数 : genFunction)



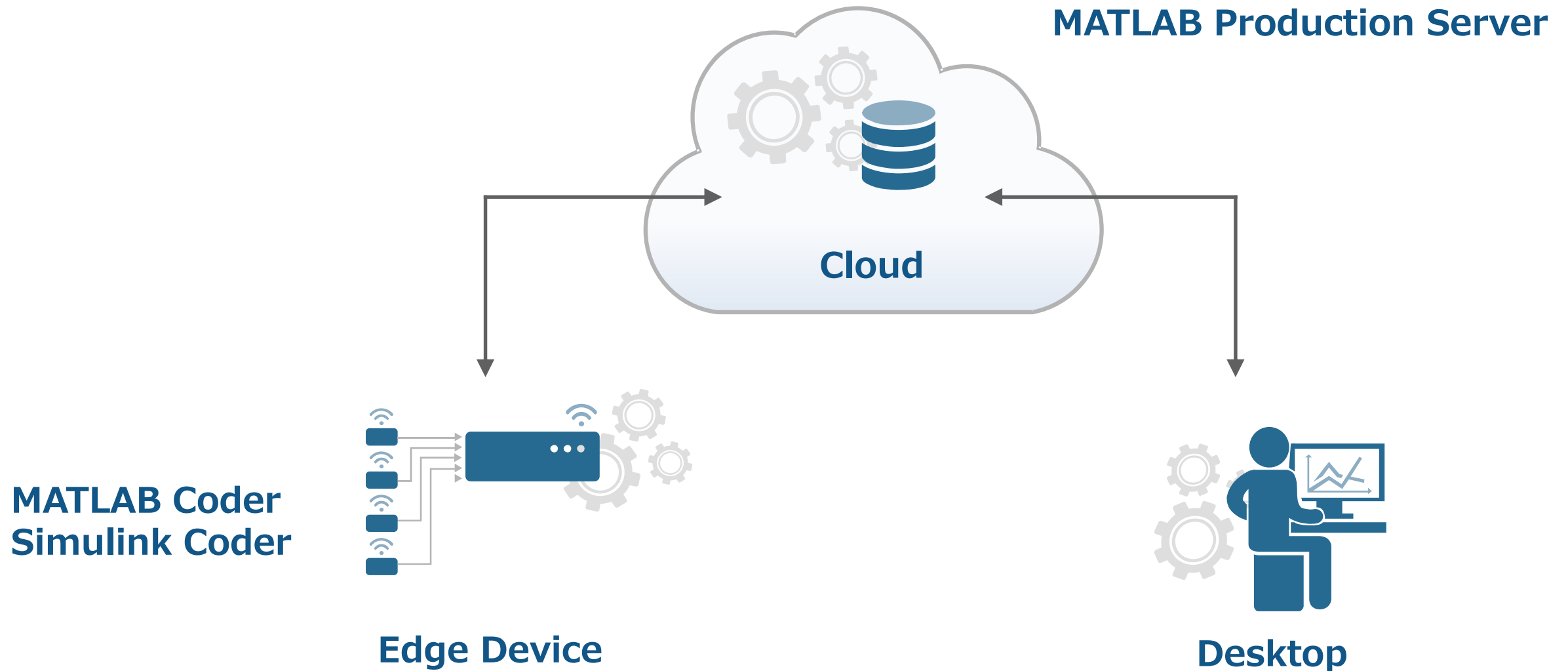
```
1 function [Y,Xf,Af] = neural_function(X,~,~)
2 %NEURAL_FUNCTION neural network simulation function.
3 %
4 % Generated by Neural Network Toolbox function genFunction, 08-Jul-2016 15:06:58.
5 %
6 % [Y] = neural_function(X,~,~) takes these arguments:
7 %
8 % X = 1xTS cell, 1 inputs over TS timesteps
9 % Each X{1,ts} = 4xQ matrix, input #1 at timestep ts.
10 %
11 % and returns:
12 % Y = 1xTS cell of 1 outputs over TS timesteps.
13 % Each Y{1,ts} = 3xQ matrix, output #1 at timestep ts.
```

MATLAB関数

MATLAB Production Server™



MATLAB/Simulink によるセンサーデータ解析システムの構築



Questions?