

# at-13.全体まとめと発展

(ディープラーニングのシステムとプログラミング)  
(全12回)

<https://www.kkaneko.jp/ai/at/index.html>

金子邦彦





# 15-1. はじめに

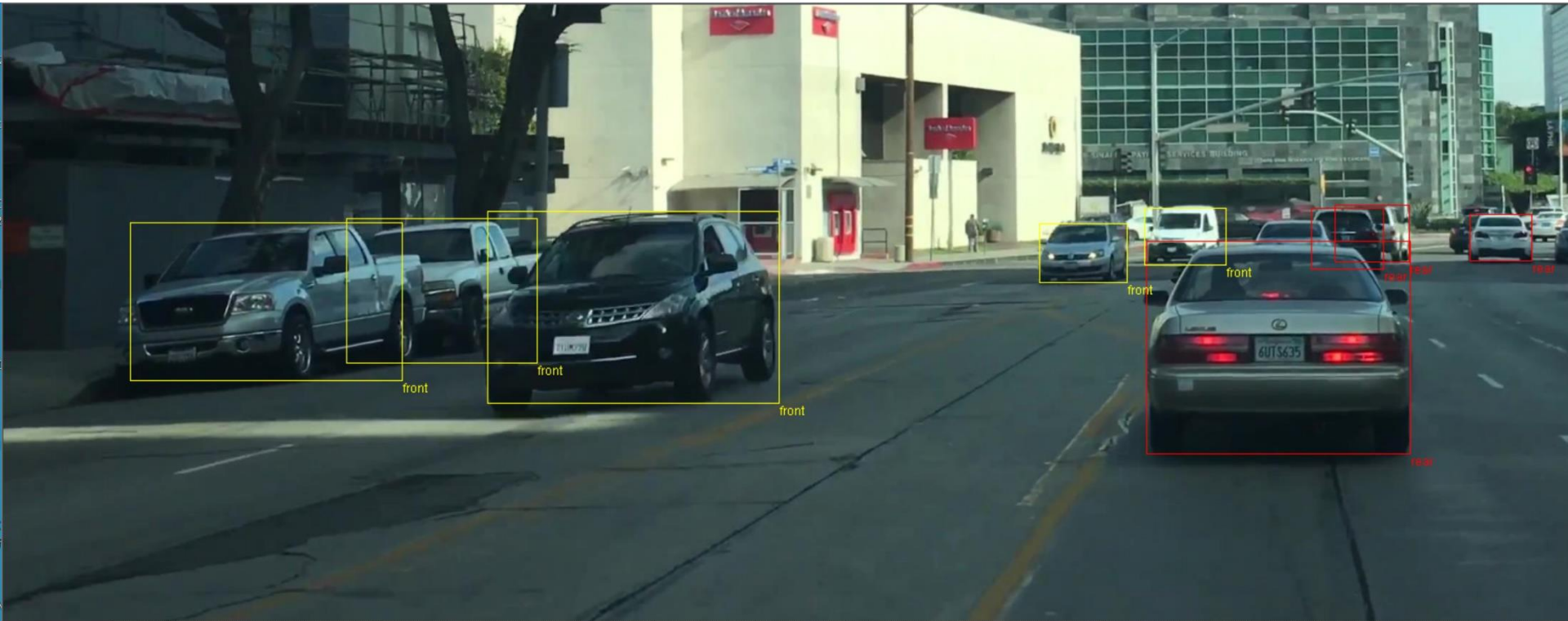
- 人工知能は、コンピュータが知的な能力を持つこと

知能：思考や判断などの能力

知識：情報を扱う能力

学習：知的な能力が上達できる能力

# 車両の発見・検知



人工知能は、車両の場所と**向き**（前なのか後ろなのか）を素早く発見できるようになってきた  
（Dlib を使用）

# 人や自転車などの，オブジェクトの発見・検知



元画像

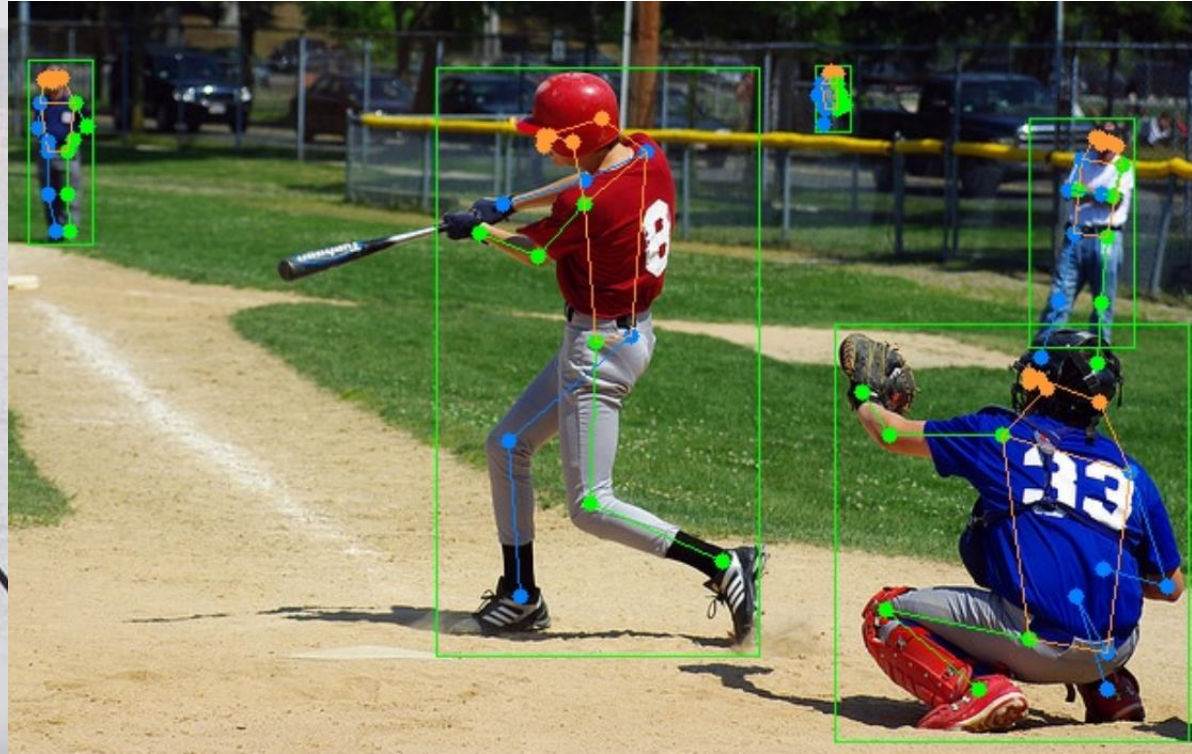
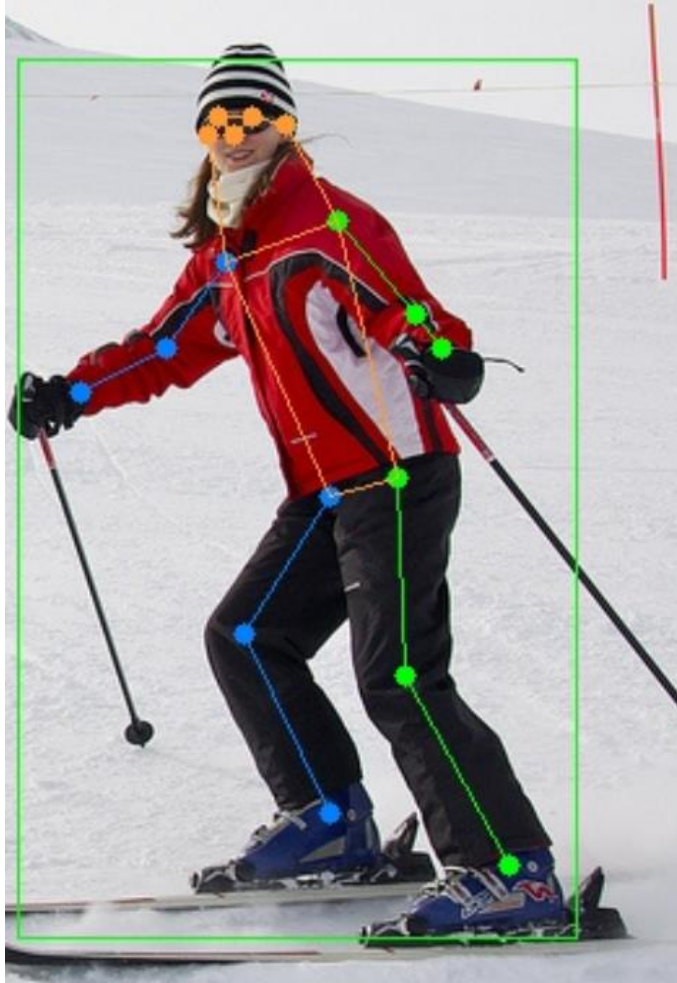


人工知能による読み取り結果  
(DeepLabv3+ を使用)

- 人間の「目」の一部機能をコンピュータで再現。  
画像の中のオブジェクトを，**人工知能**が発見・検知



# 人体の向き, ポーズの読み取り



- 写真やビデオから, 人体の姿勢を読み取り

# 翻訳を行うオンラインサービス



## Web ブラウザで動く

A screenshot of the DeepL online translation interface. At the top, there are two tabs: "テキスト" (Text) and "ドキュメント" (Document). Below the tabs, there are language selection menus. The source language is set to "日本語 - 自動検出" (Japanese - auto-detect) and the target language is "英語" (English). The input text is "白い雲と青い空が美しい" (Shiroi kumo to aoi sora ga utsukushi) and the output is "Beautiful white clouds and blue sky". There are also icons for voice input/output, a character count (11/5000), and a feedback button at the bottom right.

日本語 - 自動検出 英語 日本語 韓国語 ▼ ↔ 日本語 英語 韓国語 ▼

白い雲と青い空が美しい × Beautiful white clouds and blue sky ☆

Shiroi kumo to aoi sora ga utsukushi

11/5000 フィードバックを送信

DeepL の URL: <https://www.deepl.com/ja/translator>

- **人工知能（AI）の究極の目標は、コンピュータで人間の知能を模倣すること**
- 学習、問題解決、パターン認識、対話などを行う
- **AIは、すでに、私たちの日常生活に深く浸透している**

例： スマートフォンの音声アシスタント  
自動運転車

- **生活と社会が大きく変化**

例：

- 医療診断，個別化医療の実現
- 自動運転技術，交通事故の減少
- 言語の壁を越えたコミュニケーションの促進





## 15-2. ディープラーニング

# 人工知能

知的なITシステム

## 機械学習

データから**学習**し、知的能力を向上

## ディープラーニング

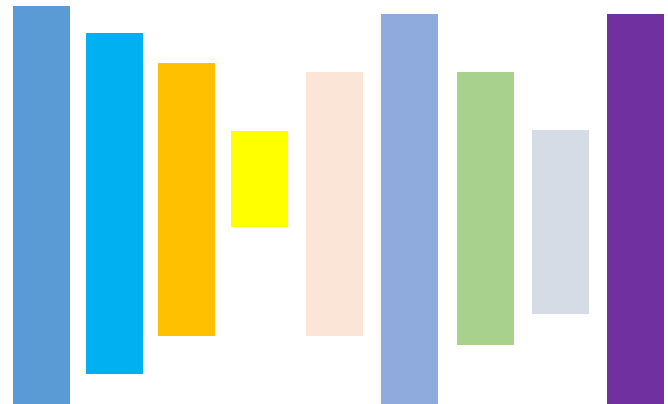
データから**学習**し、複雑なタスクを実行。**多層のニューラルネットワーク**を使用

# ディープラーニング

ディープラーニングに「ディープ」とついているのは、多層のニューラルネットワークを使用するため



層の数が少ない



層の数が多し (ディープ)

# ディープラーニングまとめ



- **ディープラーニング**は**機械学習**の一種であり、人工ニューラルネットワークを使用して**データから学習**し、**複雑なタスクを実行**する技術
- 「ディープ」の名前は、**多層のニューラルネットワーク**を使用することに由来
- ディープラーニングが広く利用される理由は、**多様なデータに適用**でき、**さまざまなタスク**で高性能を発揮するため。  
例：**画像認識**、**自然言語処理**、**音声認識**など。



# 15-3. ニューラルネットワーク

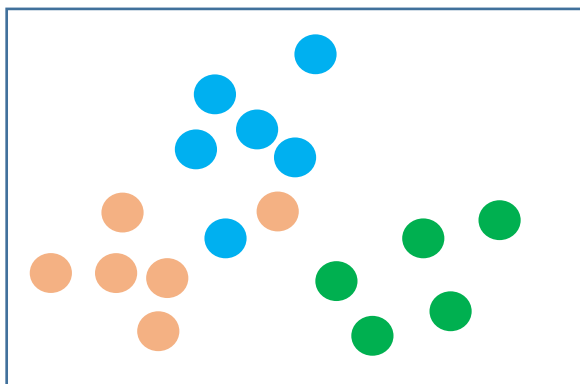
- ニューラルネットワークの究極の目標は、**人間の脳の仕組みを模倣**すること

## 《特徴》

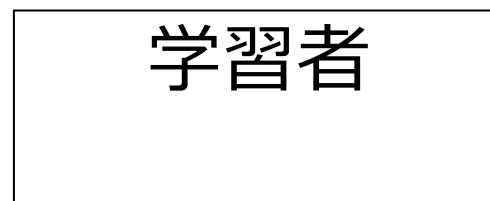
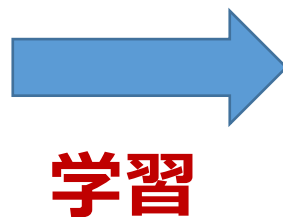
- **大量のデータ**から自動的に**パターンを学習**
- **複雑な問題を解決する能力**（画像認識、自然言語処理など）

## 機械学習は、データによる学習を行うための手法 (人工知能の一種)

### 学習用データ



3種類に分類済み



大量の学習用データを用いて学習を行う



# ニューラルネットワークの処理の原理



## 1. 入力の重みづけ:

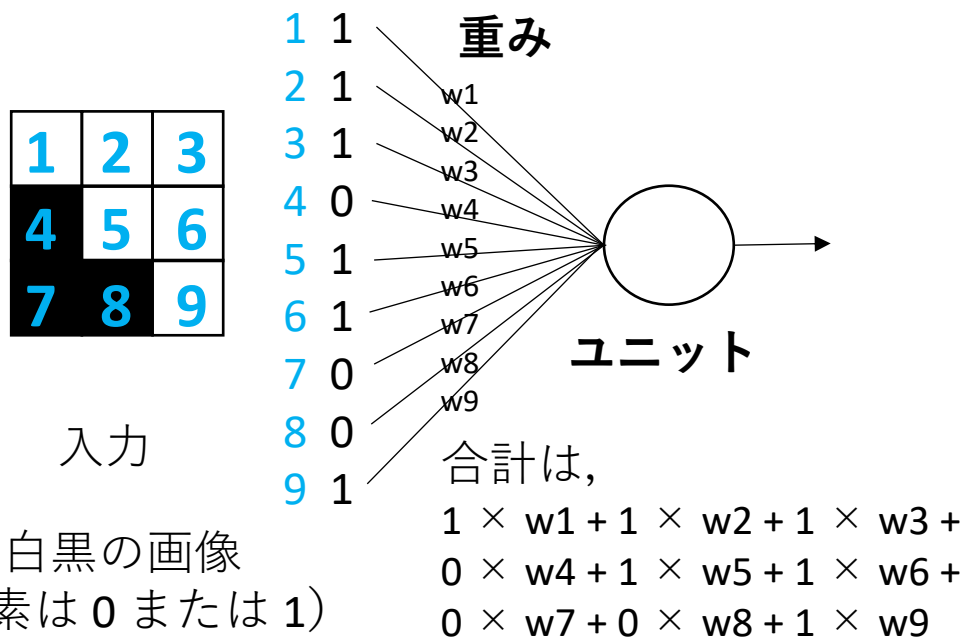
$1 \times w_1, 1 \times w_2, 1 \times w_3, 0 \times w_4, 1 \times w_5, 1 \times w_6, 0 \times w_7, 0 \times w_8, 1 \times w_9$   
( $w_1$  から  $w_9$  は重み)

## 2. 合計とバイアス:

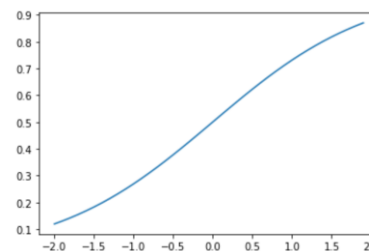
$1 \times w_1 + 1 \times w_2 + 1 \times w_3 + 0 \times w_4 + 1 \times w_5 + 1 \times w_6 + 0 \times w_7 + 0 \times w_8 + 1 \times w_9 + b$  ( $b$  はバイアス. プラスや0やマイナスの数)

## 3. 活性化関数の適用による出力値の取得

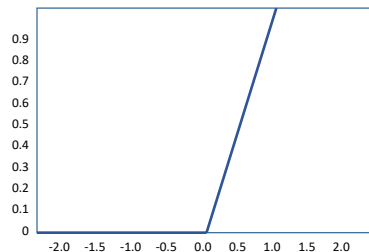
$f(1 \times w_1 + 1 \times w_2 + 1 \times w_3 + 0 \times w_4 + 1 \times w_5 + 1 \times w_6 + 0 \times w_7 + 0 \times w_8 + 1 \times w_9 + b)$  ( $f$  は活性化関数)



## 活性化関数はさまざまな種類



シグモイド



ReLU  
(2011年発表)

- **ニューラルネットワークは、単純な数学的操作（足し算、掛け算、活性化関数の適用）の組み合わせで動作**
- 単純な仕組みであるが、学習能力を持つ
- **複雑な問題を解決する能力**（画像認識、自然言語処理など）

# コンピュータの進展とニューラルネットワーク進展



年	コンピュータで扱えるユニット数の規模
2010年	100,000個
2020年	2,000,000個
2030年	50,000,000個
2040年	1,000,000,000個
2050年	20,000,000,000個

諸説あります

2055年ごろには、860億を超えるかも？

## 【過去】

人間の脳の約860億個のニューロンを模倣することは不可能

## 【現状】

GPU, クラウドコンピューティングで、コンピュータの性能向上が続く

## 【将来】

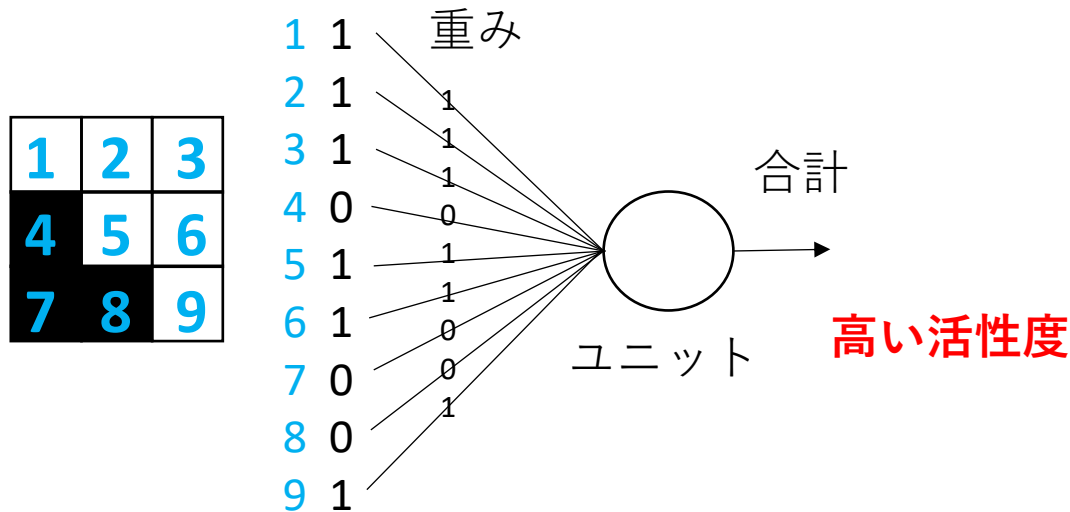
生物の脳に近いニューラルネットワークの実現へ進む



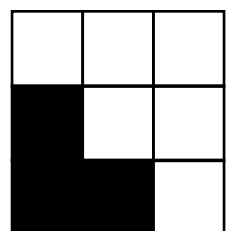
# 15-4. 最適化

- **最適化は、与えられた条件の中で最良の解を見つけ出す過程。**

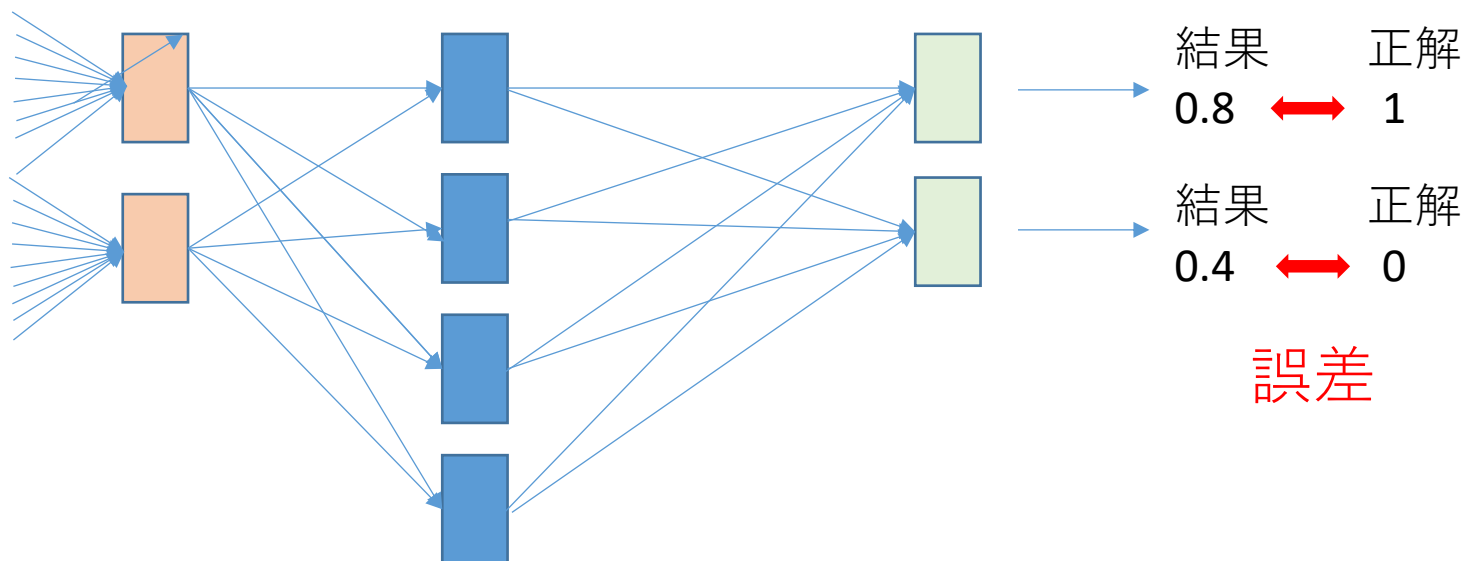
# ニューラルネットワークの学習では、結合の重みとバイアスの調整により、訓練データのパターンをより正確に認識できるようになる



# ニューラルネットワークでの教師あり学習



入力



実際に一度動作させてみて、  
誤差が少なくなるように、  
結合の重みとバイアスを調整  
= ニューラルネットワークの学習



# ニューラルネットワークでの最適化



- **ニューラルネットワークの学習過程自体が、一種の最適化**になっている。
- ニューラルネットワークの膨大なニューロンを調整して、**最も良い結果を出すための「結合の重み」と「バイアス」**を見つけ出す



# 15-5. 汎化

# 汎化



- **汎化は、学習したことを未知の状況に適用する能力になる**



## 訓練データと未知のデータ

訓練データ	
入力	正解
9	5 0 0
1 1	5 0 0
1 2	1 0 0 0
1 4	1 0 0 0

入力	予測結果
<b>7</b>	<b>5 0 0</b>
<b>8</b>	<b>5 0 0</b>
9	5 0 0
<b>1 0</b>	<b>5 0 0</b>
1 1	5 0 0
1 2	1 0 0 0
<b>1 3</b>	<b>1 0 0 0</b>
1 4	1 0 0 0
<b>1 5</b>	<b>1 0 0 0</b>
<b>1 6</b>	<b>1 0 0 0</b>

訓練データに基づいて学習し，未知のデータに対しても適切に処理

# ① 一般のプログラミング

- ・プログラムは人間が作成してテストし，調整する。

データ  
(入力)



プログラム

コンピュータ



処理結果

# ② 機械学習

- ・データを利用して知的能力を向上させる

データ  
(入力)



プログラム

コンピュータ



処理結果

訓練データ



## 《汎化のメリット》

- 新しい状況に対応できる
- 人間の学習能力を模倣できる

例：手書き数字認識AIは、学習時にで見たことのない人の字でも正しく認識できる。

⇒ 数字の本質的な特徴を学習。新しい状況に対応。

# 15-6. 時系列データと予測



# 時系列データ

時系列データは、時間の経過に伴って順序付けられたデータの並び

例

2023年12月3日の気温は15度

2023年12月4日の気温は13度

...

# 時系列データの例：太陽黒点観測データ

11314	1848	12	23	1848.977	353	23.8	1
11315	1848	12	24	1848.980	240	19.6	1
11316	1848	12	25	1848.982	275	21.0	1
11317	1848	12	26	1848.985	352	23.8	1
11318	1848	12	27	1848.988	268	20.8	1
11319	1848	12	28	1848.990	285	21.4	1
11320	1848	12	29	1848.993	343	23.5	1
11321	1848	12	30	1848.996	340	23.4	1
11322	1848	12	31	1848.999	238	19.6	1
11323	1849	1	1	1849.001	287	20.9	1
	year	month	day	dec_year	sn_value	sn_error	obs_num
72855	2017	6	21	2017.470	35	1.0	41
72856	2017	6	22	2017.473	24	0.8	39
72857	2017	6	23	2017.475	23	0.9	40
72858	2017	6	24	2017.478	26	2.3	15
72859	2017	6	25	2017.481	17	1.0	18
72860	2017	6	26	2017.484	21	1.1	25
72861	2017	6	27	2017.486	19	1.2	36
72862	2017	6	28	2017.489	17	1.1	22
72863	2017	6	29	2017.492	12	0.5	25
72864	2017	6	30	2017.495	11	0.5	30

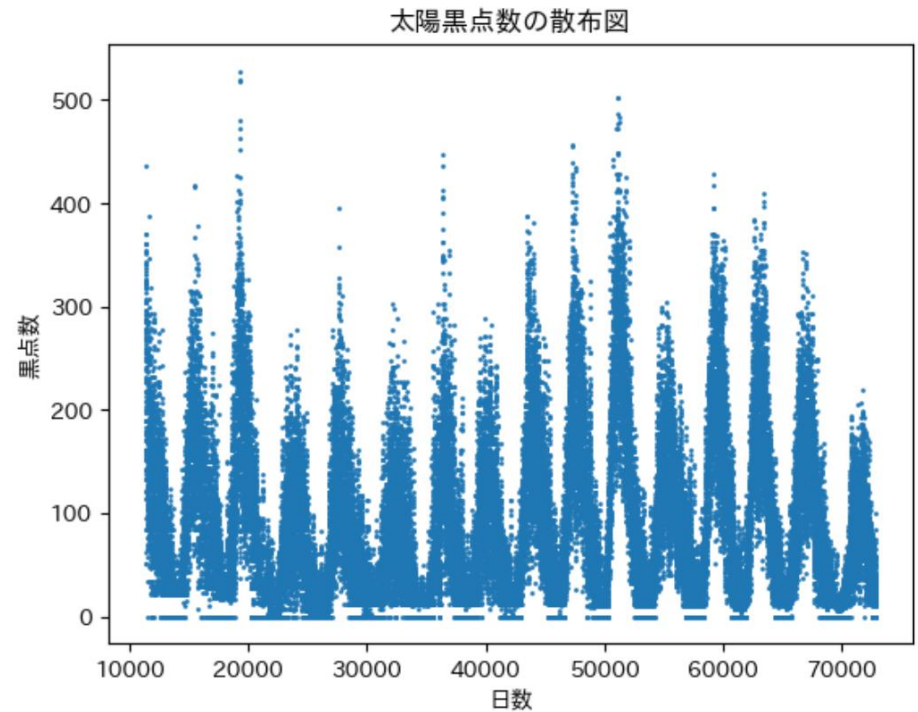


年月日の値 黒点数

黒点数の観測データ

1848年12月23日～2017年6月30日の  
毎日の黒点数データが公開されている

[https://data.heatonresearch.com/data/t81-558/SN\\_d\\_tot\\_V2.0.csv](https://data.heatonresearch.com/data/t81-558/SN_d_tot_V2.0.csv)

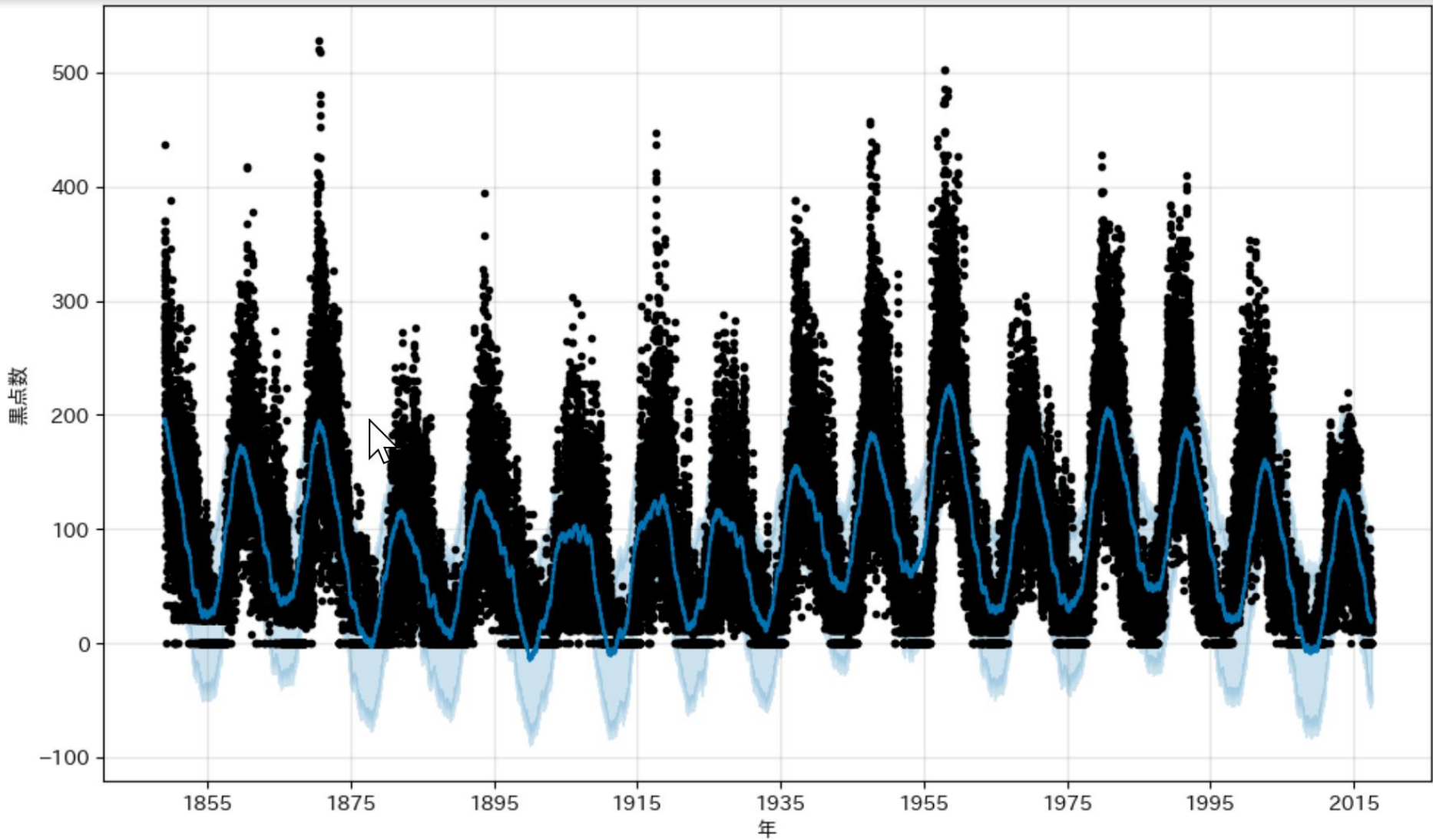


散布図でプロット

# 時系列データの特徴

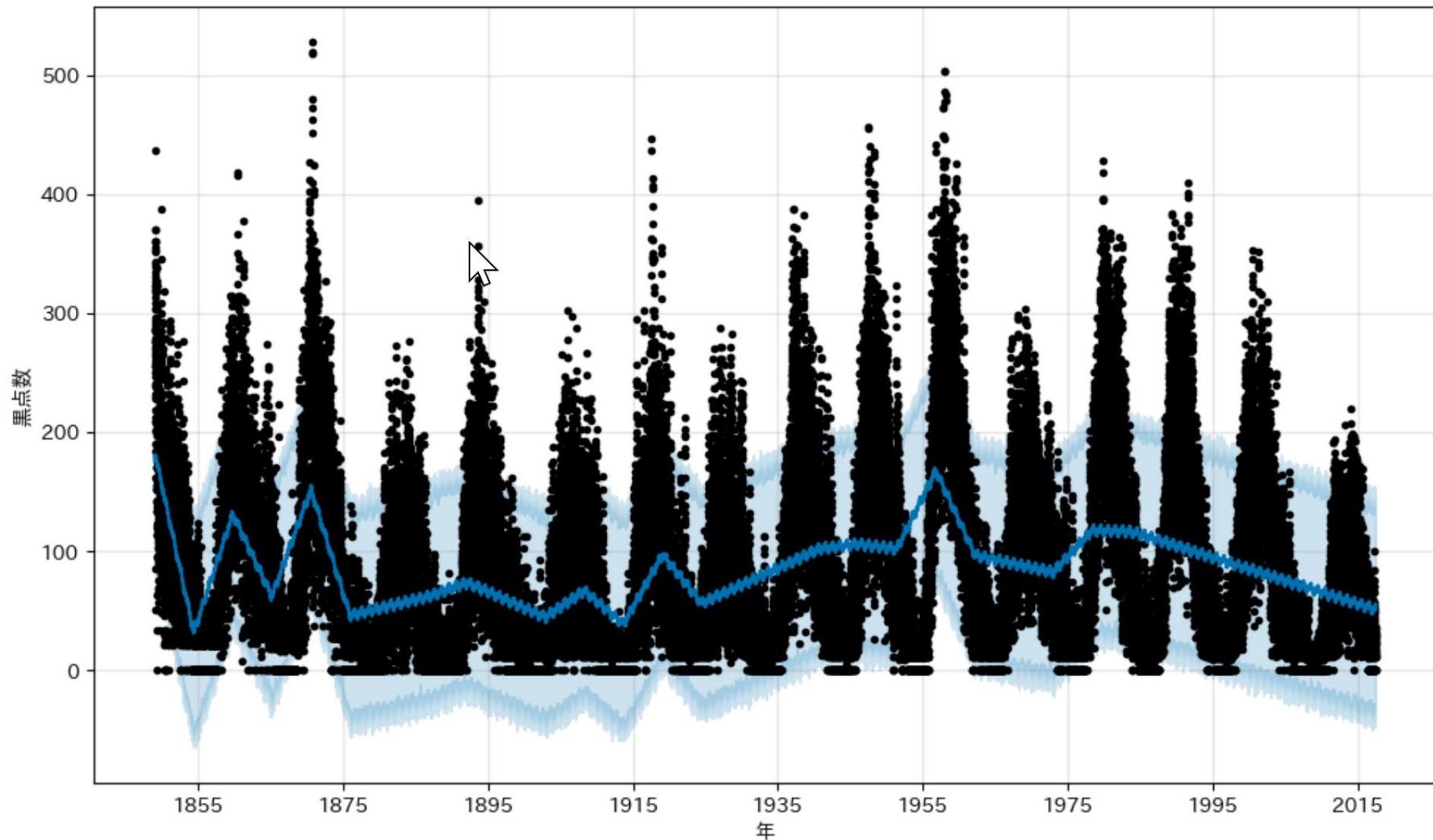
- **周期性**：日、週、月、季節などの**定期的な間隔で繰り返されるパターン**  
例：自然現象、社会的活動
- **トレンド**：時間の経過による**増加、減少、一定レベルの維持などの方向性**
- **特定のイベントや時期**（例えば正月、学校の学期開始時期など）**との関連性**

# 1 1 年周期の周期性の分析の例 (太陽黒点データ) Prophet を使用



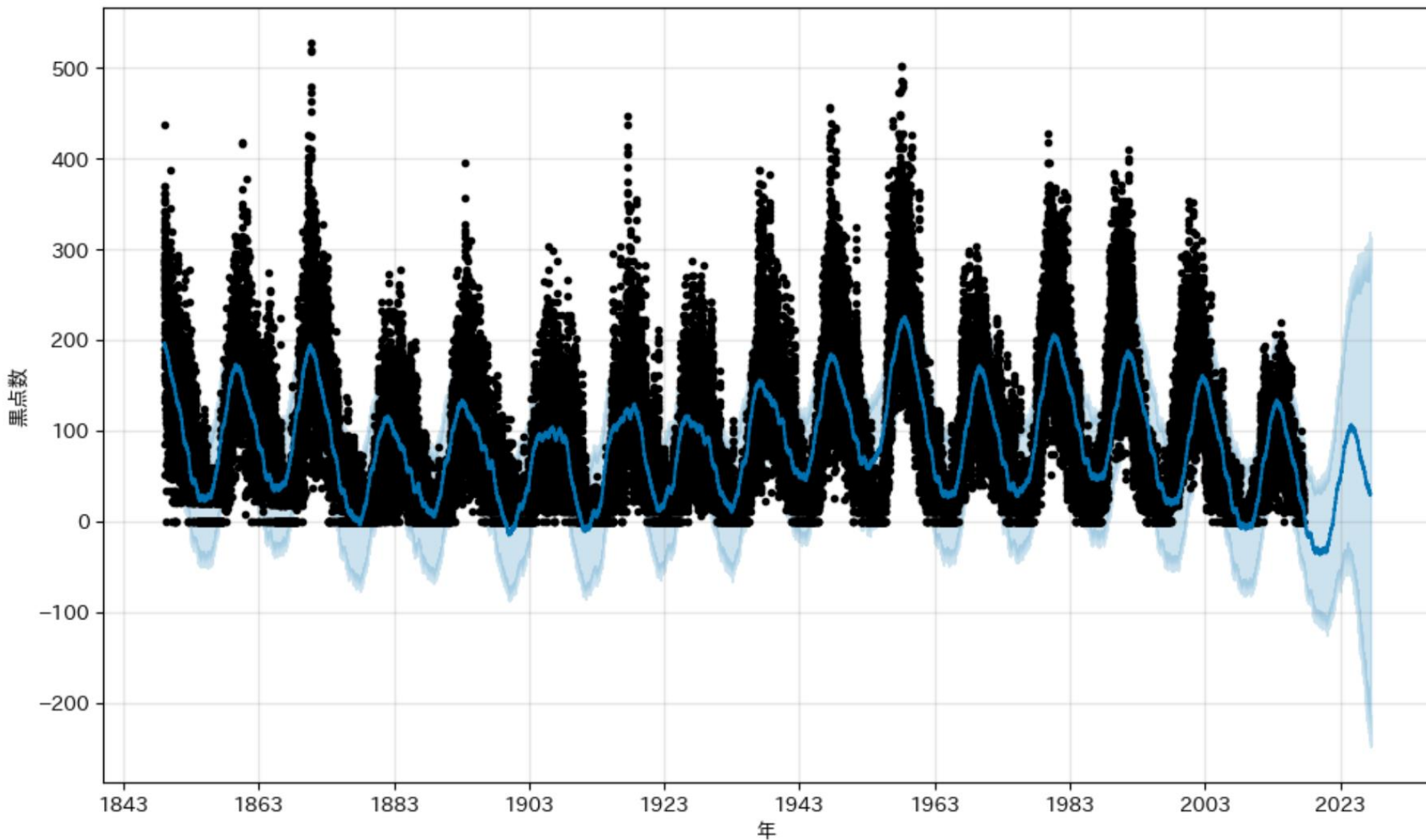
# トレンドの識別の例 (太陽黒点データ) Prophet を使用

太陽の黒点数の分析



# 将来予測の例 (太陽黒点データ) Prophet を使用

太陽の黒点数の予測 (10年先まで)



# Prophet



- Prophet は Python のライブラリ
- **時系列データ**に対して、**周期性、トレンド、特定のイベントや時期との関連性を分析**する機能を持つ
- 統計手法を基礎とする。線形モデルと非線形モデルを組み合わせた回帰モデル。
- ディープラーニングではない

# リカレントニューラルネットワーク (RNN)



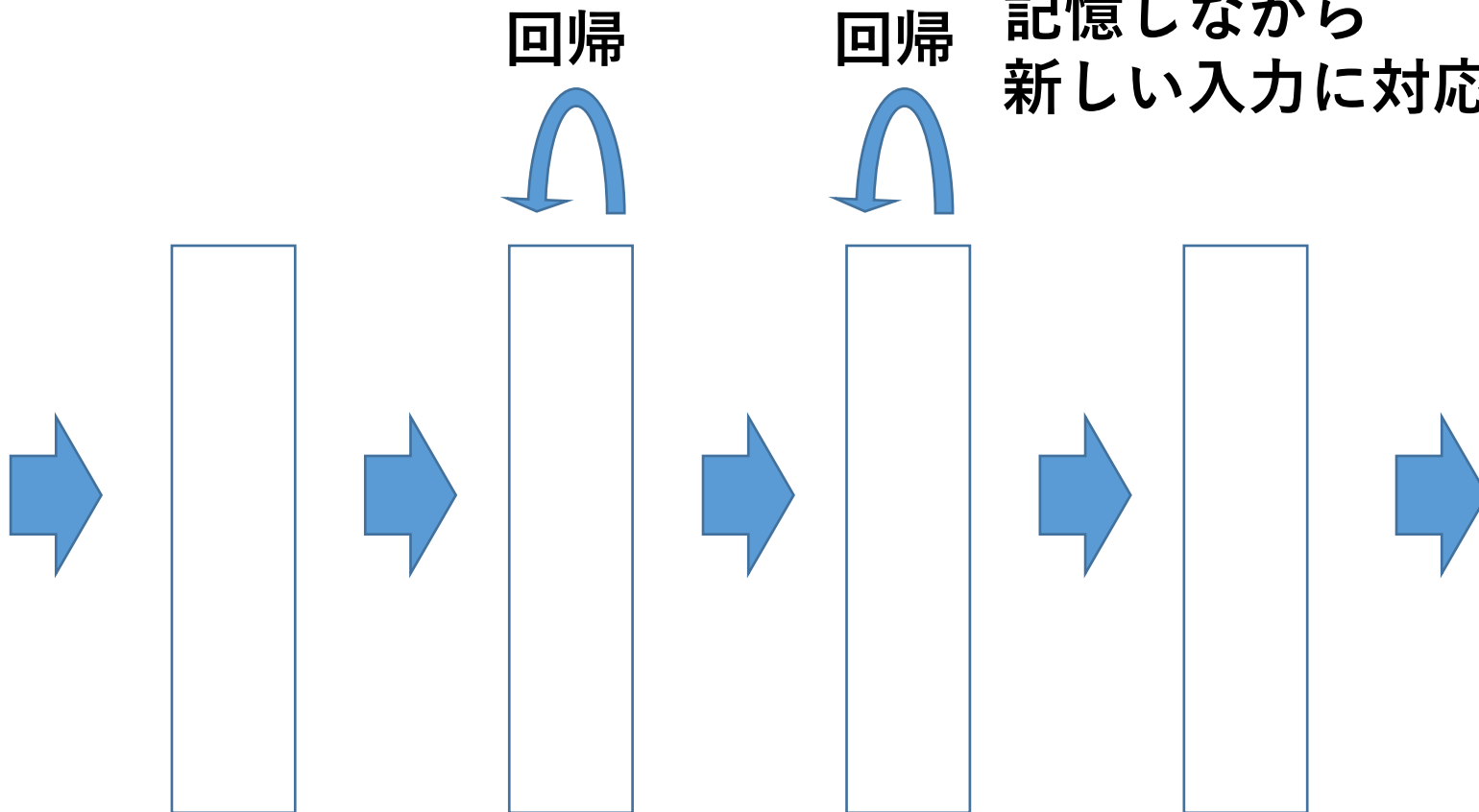
- **リカレントニューラルネットワーク (RNN) は、時系列データや、その他データの並びに適したニューラルネットワークの一種**
- **回帰のしくみが導入されている**
- **前回の実行時での結果を記憶しながら新しい入力に対応**



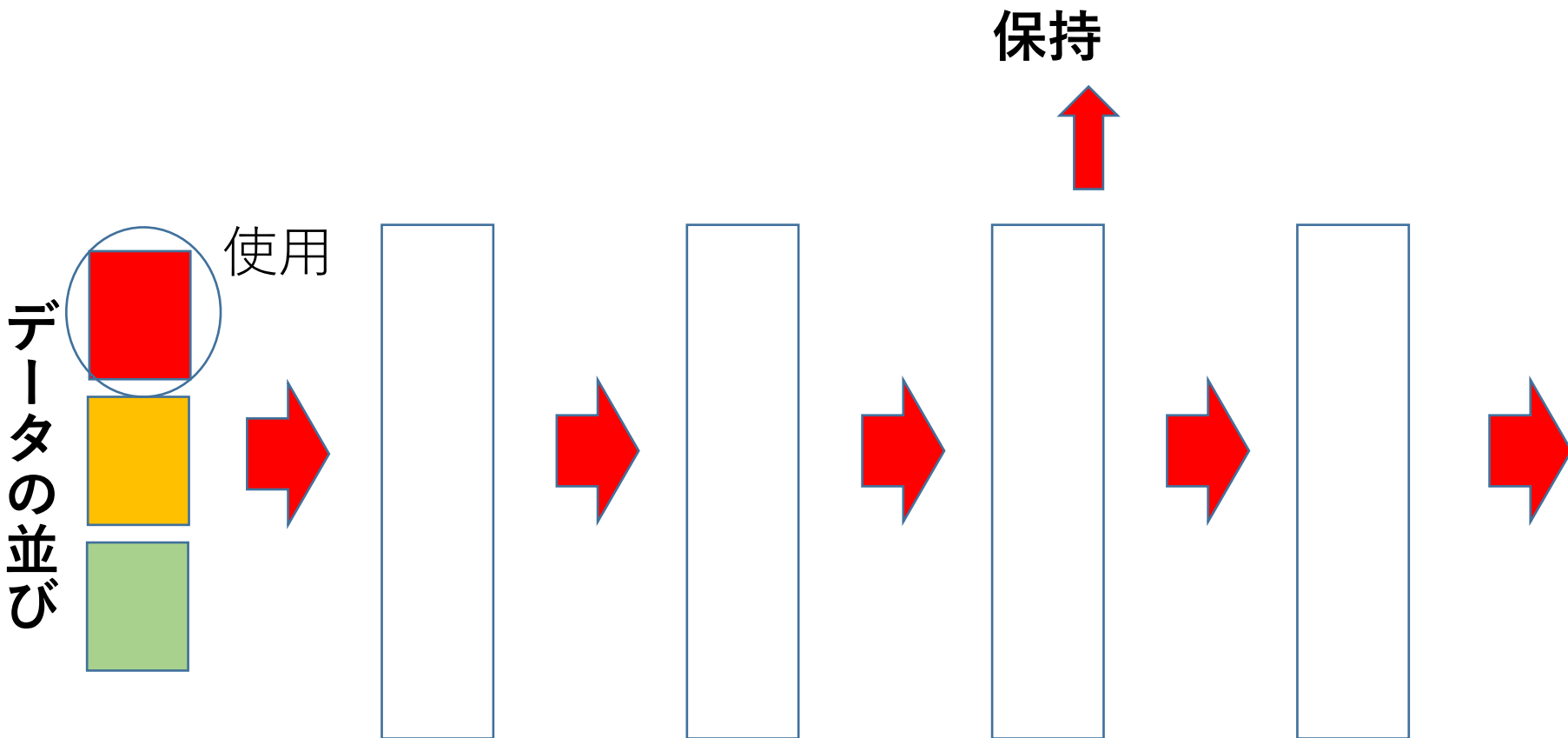
# リカレントニューラルネットワーク

## 回帰のしくみ

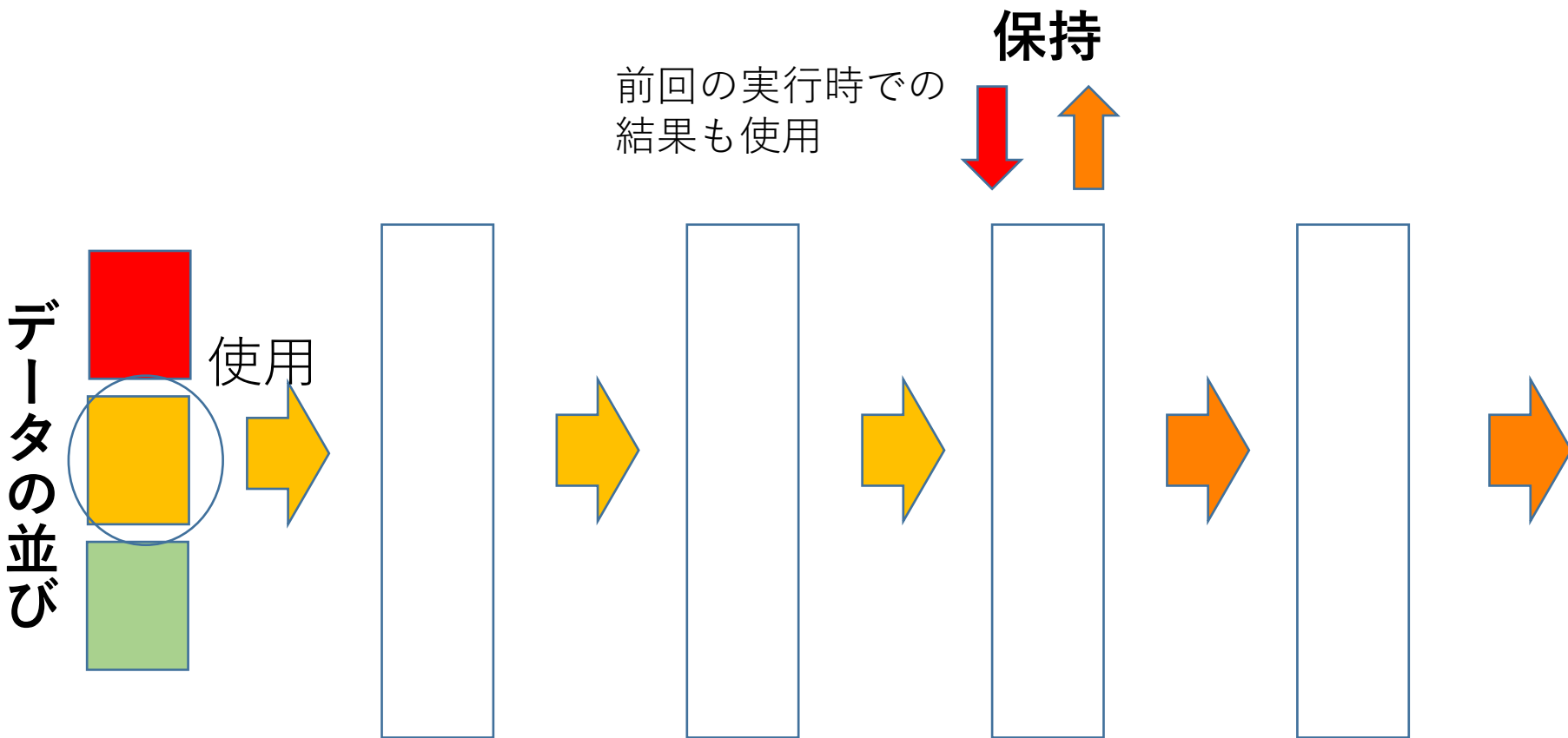
前回の実行時での結果を  
記憶しながら  
新しい入力に対応



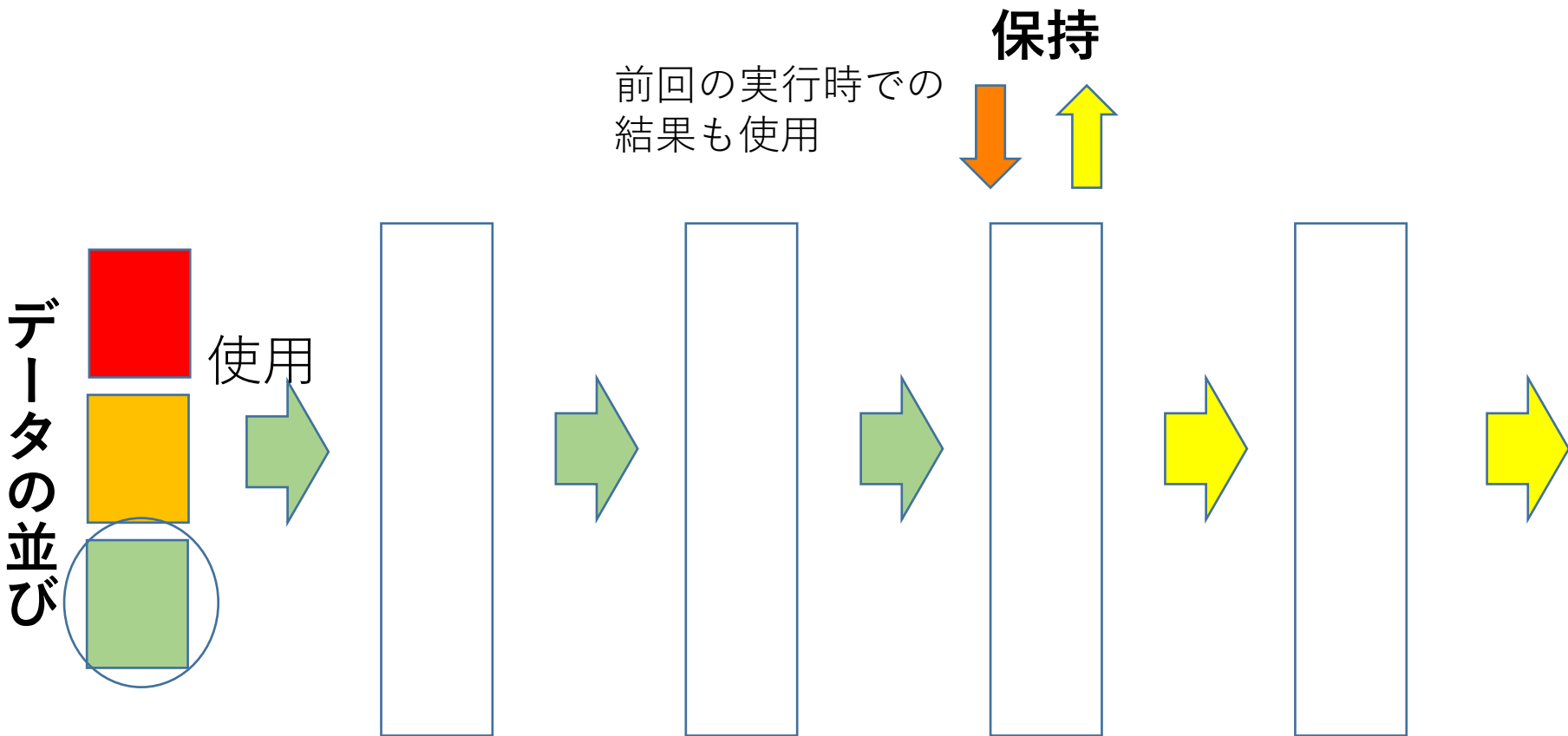
# リカレントニューラルネットワークの動作イメージ ①



# リカレントニューラルネットワークの動作イメージ ②



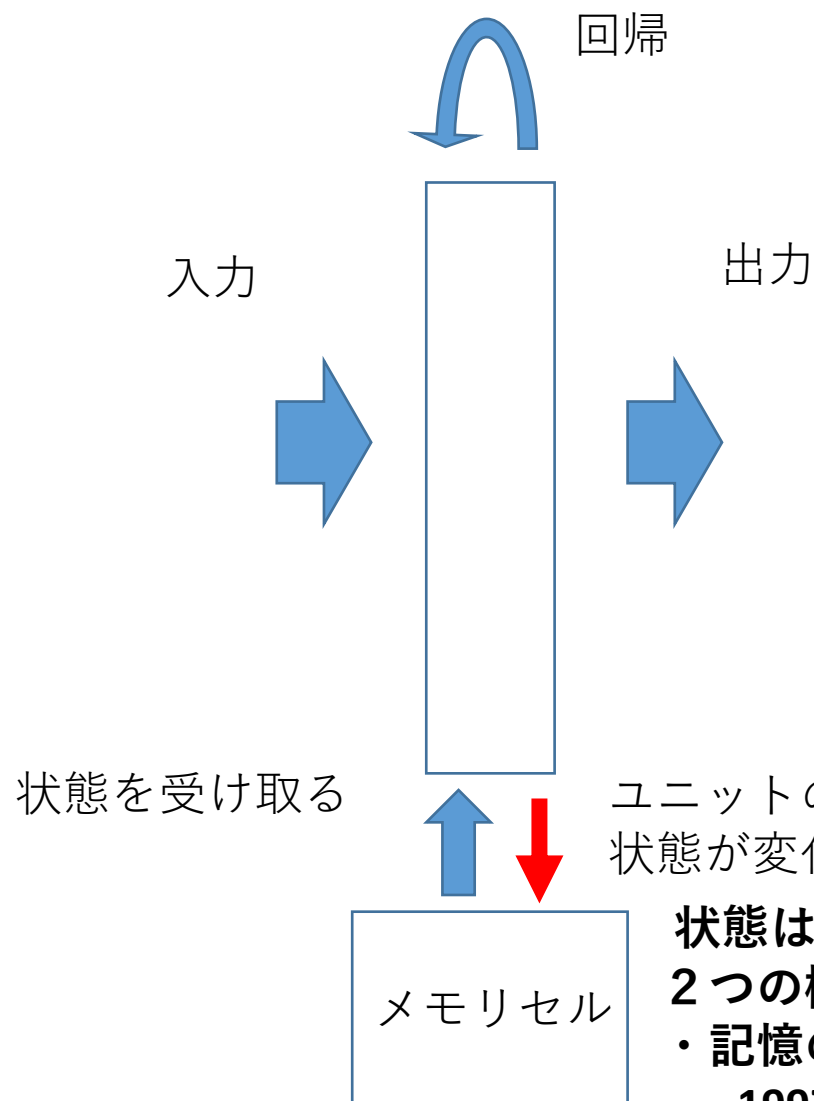
# リカレントニューラルネットワークの動作イメージ ③



# LSTM の特質

- リカレントニューラルネットワーク（RNN）の一種。1997年
- 「**長期的な依存関係の学習が困難**」という、RNNの弱点を改良
- LSTM は、長期の記憶の保持能力を持ち、長期的な依存関係の学習を可能とする。複雑なデータを扱えるように。

# LSTM の仕組み

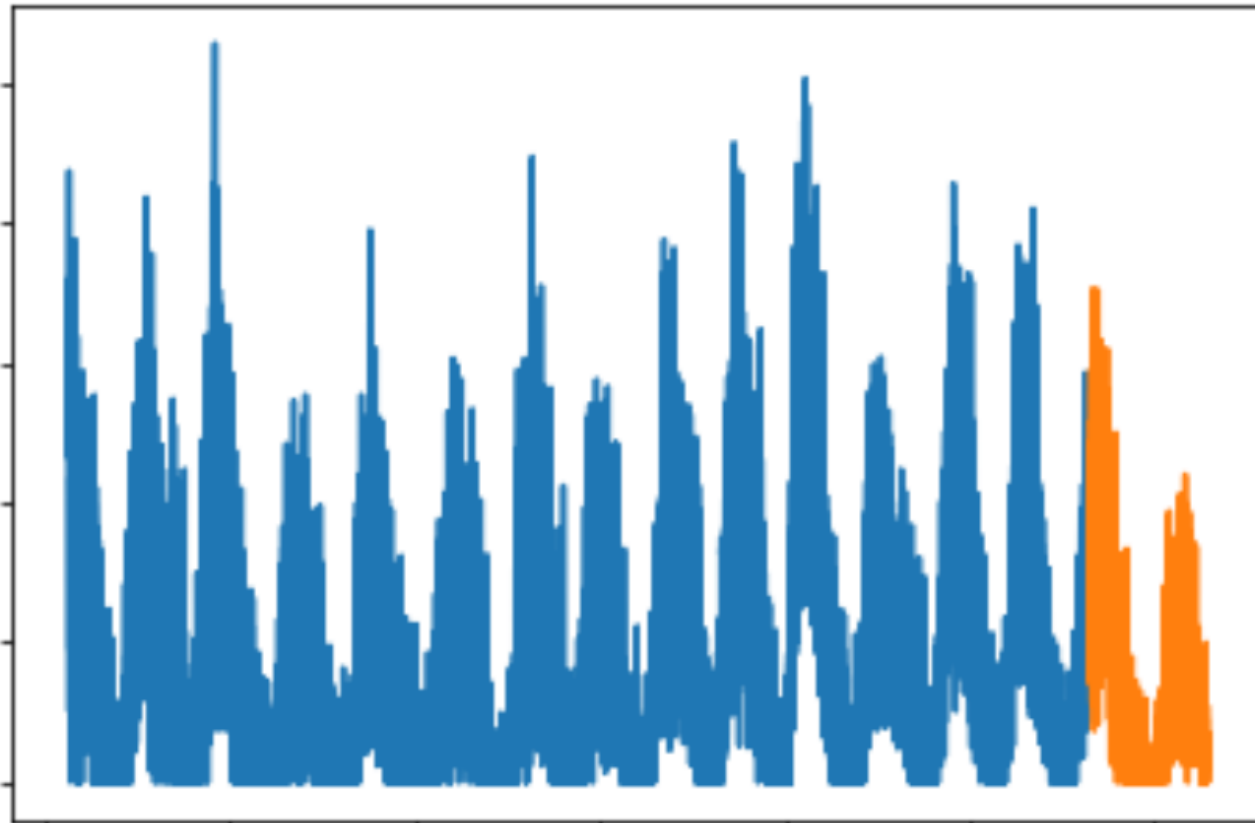


**LSTM のメモリセルは、長期の記憶の保持を可能とする**

状態は「メモリセル」に記憶されている  
2つの機能

- ・ 記憶の持続 (constant error carousel)  
1997年発表
- ・ 記憶の忘却 (forget gate)  
1999年発表

# LSTM を用いた予測の例



太陽の黒点数の変化

予測

**1848年～1999年のデータ**を用いて、**2000年以降を予測**  
(ディープニューラルネットワークによる予測)



# 15-7. 自然言語処理と チャットボット



# 自然言語処理

**自然言語処理は、人間が普段使う言語（日本語、英語など）をコンピュータが理解したり、生成する技術**

【自然言語処理のさまざまな応用】

**情報検索、AIとの対話、AIへの指示、プログラミング支援、人間の指示による文書の作成や推敲、翻訳、要約**

- **用途**：翻訳、校正、リサーチ支援、要約、プログラミング支援、自学自習支援、顧客サービス、エンターテインメント、日常業務サポートなどさまざま
- **期待される効果**：サービス品質の向上、AIと人間の共同、人間では気づきにくい過ちの発見、多言語対応など
- **利用上の注意点**
  - 不正確な情報を提供することがある。
  - 著作権に違反するコンテンツを生成するリスク。
  - プライバシーを侵害する情報をオンラインサービスに与えないように注意。
  - 大学の課題での丸写しは不適切。

# チャットボットのベストプラクティス



- **プロンプトは、チャットロボットへの質問や要求を、明確かつ具体的に**
- **以前の AI の対話に対する追加要求（明確化、追加情報の要求など）も可能**

## 適切な追加データを与える

- **追加データ（関連情報、事例、データなど）をプロンプトとして与えることで、回答の改善を行う**

## 回答の根拠の確認

- **AI の回答について、正確性や根拠を確認**
- **チャットロボットとインターネット検索を統合したサービスは、回答の根拠の確認をサポート**

# 演習 1

## チャットボット

ページ49, 50

### 【トピックス】

- 生成AIに作業を頼む
- 生成AIからアイデアや知識を引き出す

# 演習 1



次は、チャットボットのChatGPT 3.5

<https://chatgpt.com/>

注意点：

- 秘密にしたい情報を投稿してはいけません.
- AIの回答は完璧に正確というわけではありません.



# 15-8. 画像理解

# 画像理解の主な種類



## ① 画像分類

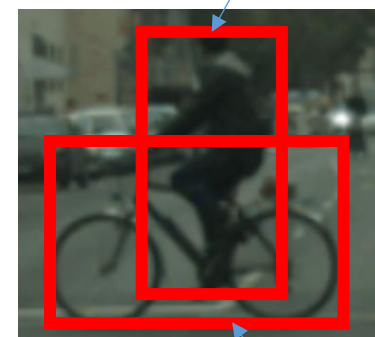
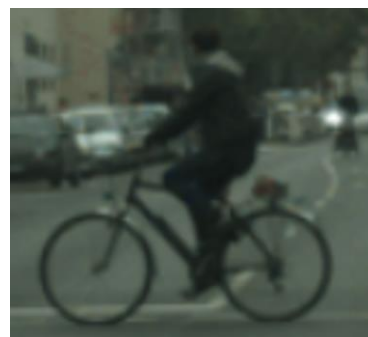
「何があるか」を理解



person  
bicycle

## ② 物体検出

場所と大きさも理解

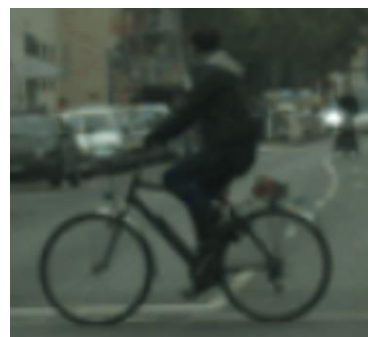


person

bicycle

## ③ セグメンテーション

画素単位で理解



# ① 画像分類

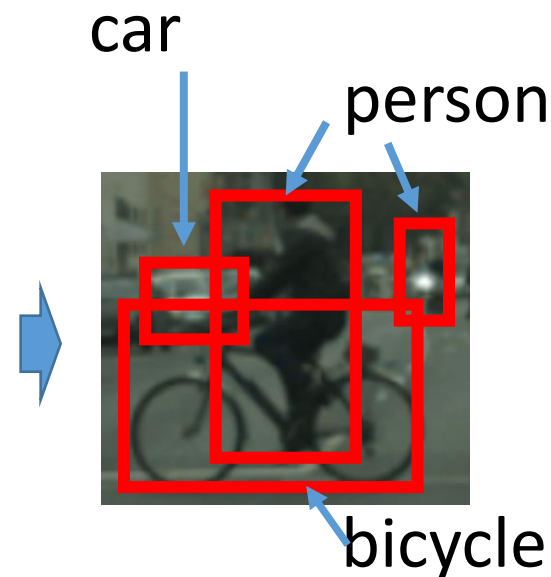
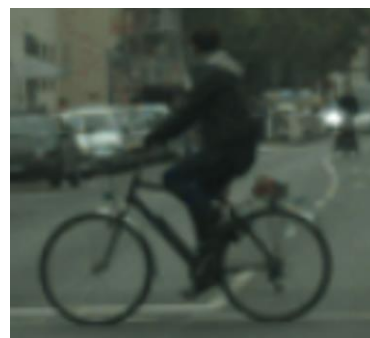


```
Score 0.9827020168304443, Label lab_coat  
Score 0.0030872616916894913, Label syringe  
Score 0.0024311079178005457, Label beaker  
Score 0.0016609227750450373, Label stethoscope  
Score 0.00037950885598547757, Label plate
```

**画像分類の結果は、ラベルと確率**  
※ 5つの候補 (top 5) が表示されている



## ② 物体検出



バウンディングボックス,  
ラベルを得る

**バウンディングボックス**は、  
物体を囲む最小のボックス（四角形）

### ③ セグメンテーション



物体の形を画素単位で抜き出し



ラベルを得ることもできる

# 画像の畳み込み

Input

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

元画像 (5 × 5 マス)

Filter / Kernel

1	0	1
1	1	1
0	0	1

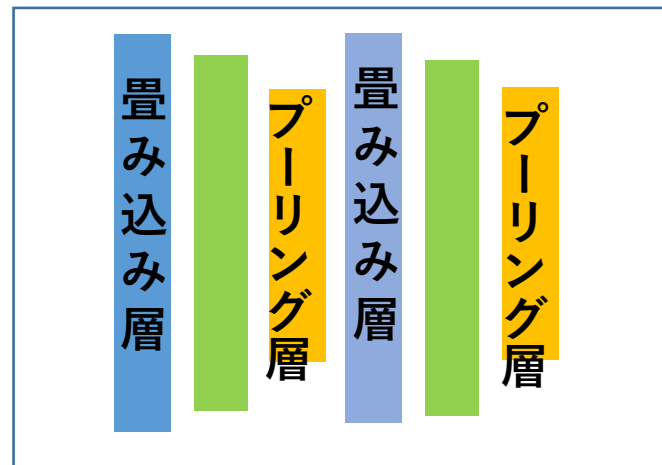
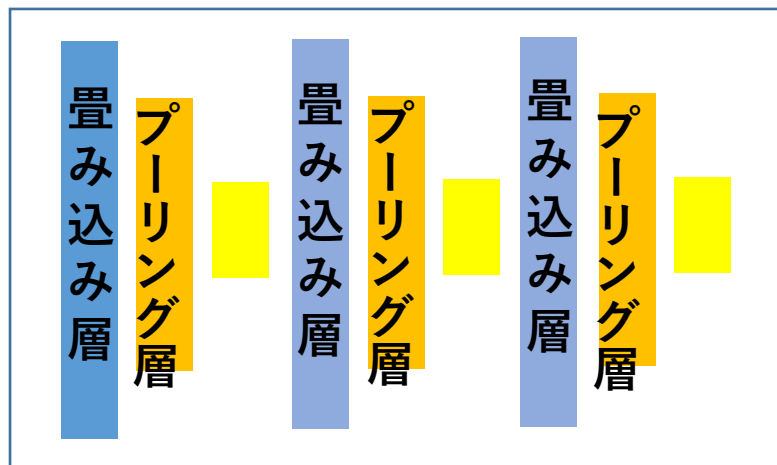
カーネル (3 × 3 マス)

# 畳み込みニューラルネットワーク (CNN)



- **畳み込みニューラルネットワーク (CNN)** は**画像理解**や**画像の分析**に特化した**ディープラーニング**の一種。
- CNNは主に**畳み込み層**、**プーリング層**、**全結合層**の3種類  
(注) これら3種類以外のもさまざまある
- **畳み込み層**：画像の局所的な特徴をとらえる役割。**特徴**は、画像内の**顕著なパターン**や**属性** (例：エッジ、テクスチャ)
- **プーリング層**：特徴マップの**サイズ**を縮小。**過学習**を防止。**計算効率**を向上
- **全結合層**：畳み込み層とプーリング層を通過した後の特徴を基に、**画像の分類**や**回帰**を行う

# 畳み込みニューラルネットワーク (CNN) の例



さまざまなバリエーション

# 畳み込みニューラルネットワークでのパターン認識

「畳み込みニューラルネットワークの利用により、さまざまなレベルのパターンを抽出・認識できるようになる」という考える場合も

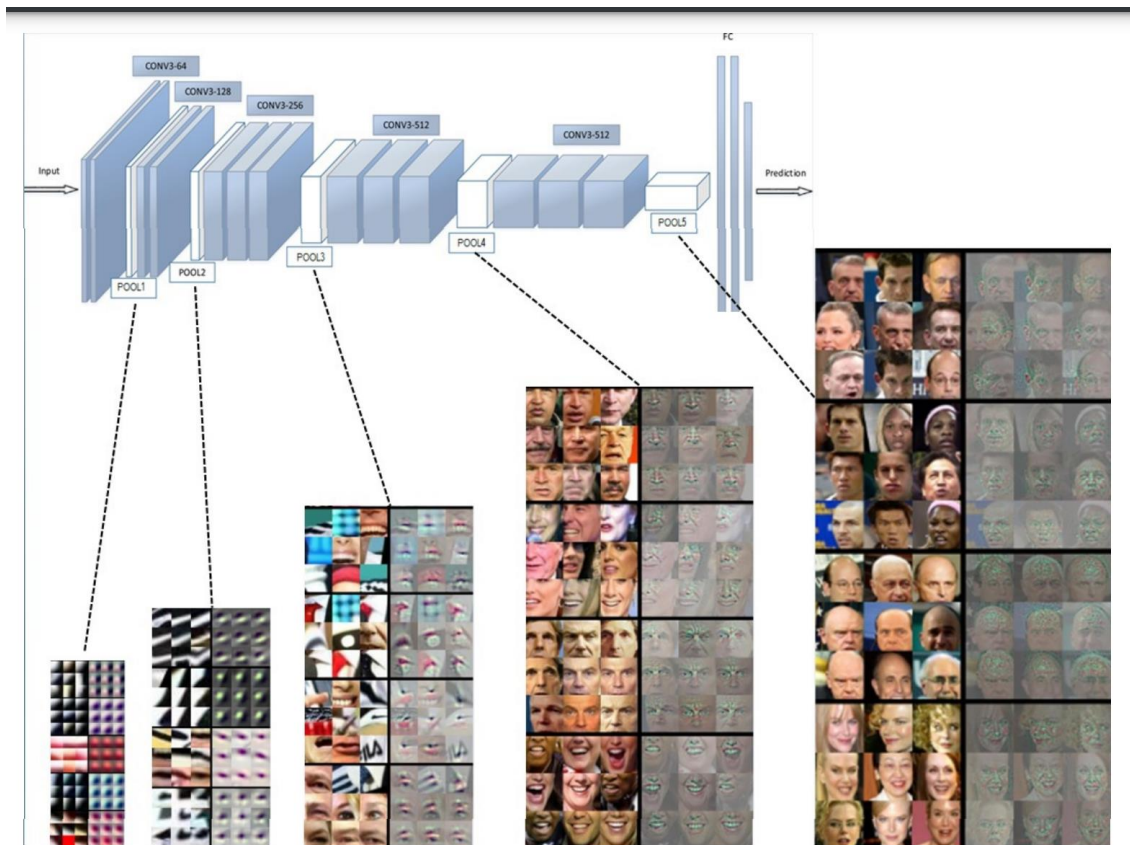


Fig. 2

Mei Wang, Weihong Deng,

Deep Face Recognition: A Survey, arXiv:1804.06655, 2018.



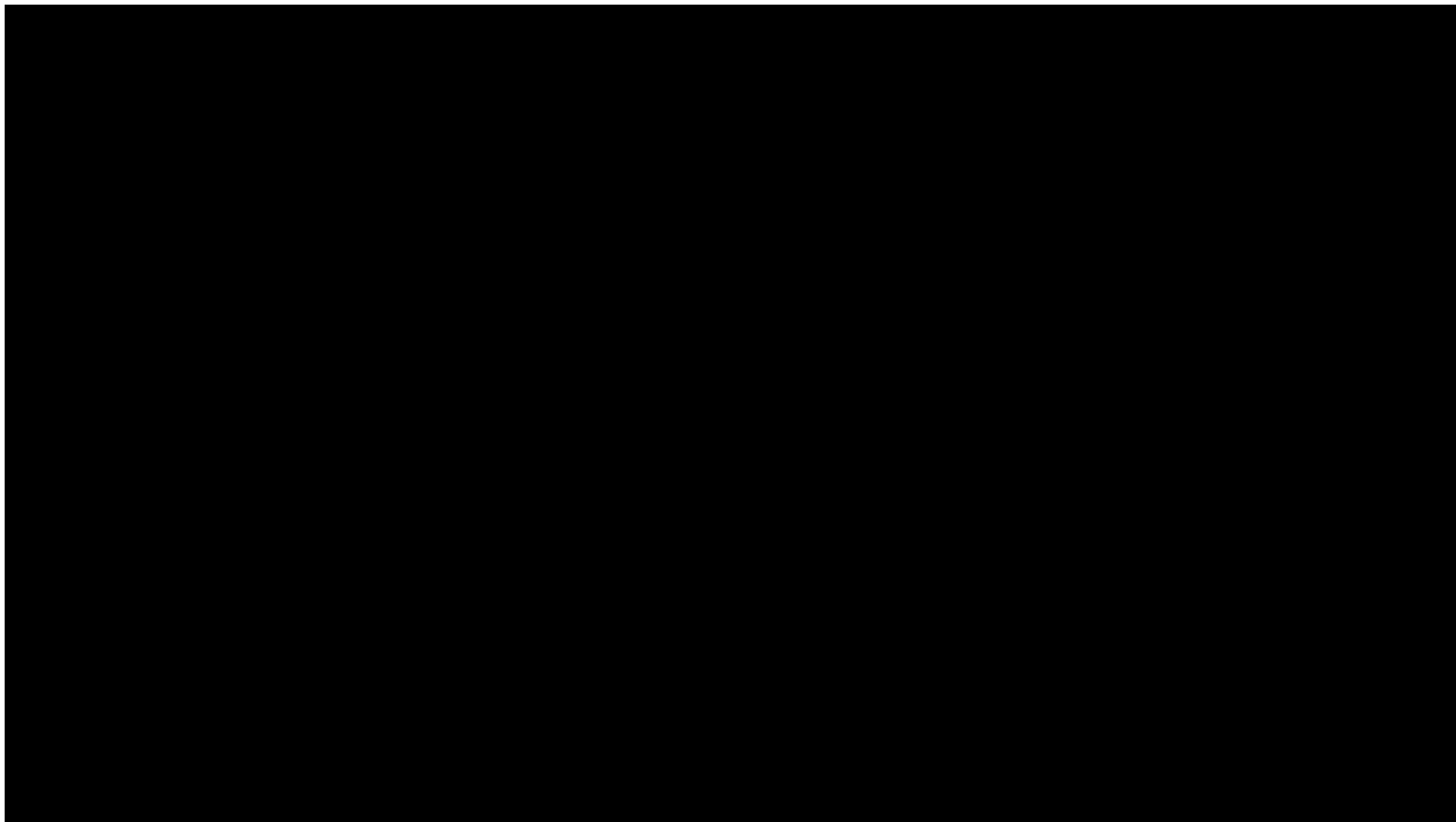


# 姿勢推定のビデオの例





# 姿勢推定のビデオの例



# 姿勢推定でのキーポイントの検出

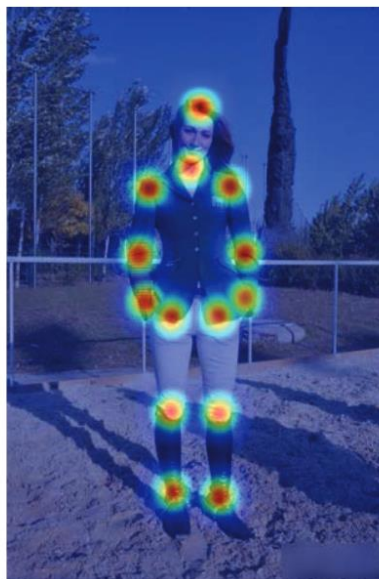
元画像

①部位の  
位置推定

②同一人物の  
キーポイント特定



(a)



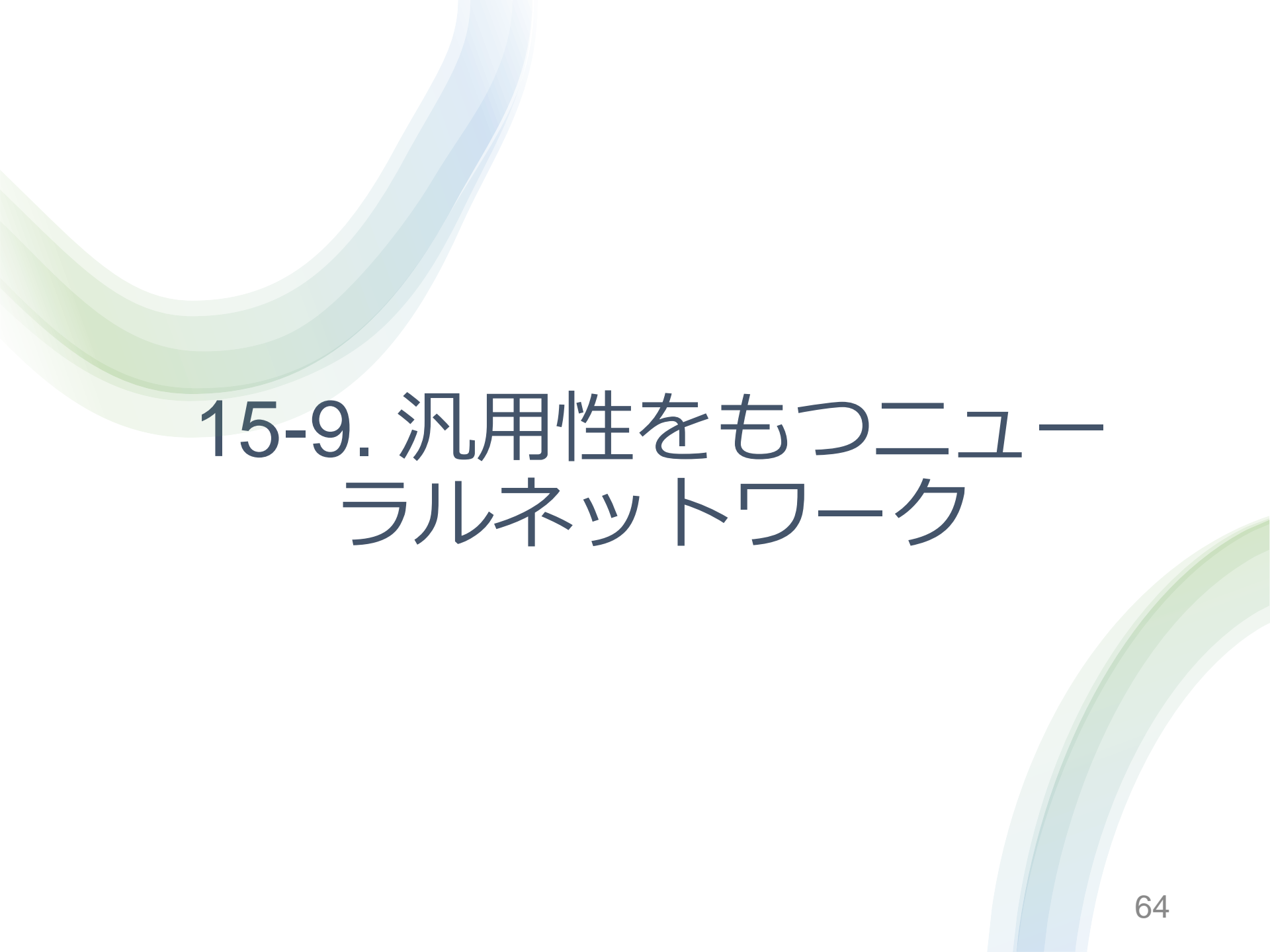
(b)



(c)

い

**Fig. 6 An example of heatmap-based single-person pipeline with heatmap. (a) Original image, (b) heatmap generated by estimator, and (c) detection result.**



# 15-9. 汎用性をもつニューラルネットワーク

## 従来の学習方法の特徴

- **学習時に使用されたラベル**（例：人間、自動車、信号機）  
**の範囲内でのみ識別・分類が可能**

## NoShot 学習の特徴

- 広範な出力能力
- **学習時のラベルにはない新しいラベルに対しても、識別・分類が可能。**

# NoShot のセグメンテーションの例



Text Prompt

hard object

画像と英語のプロンプトを  
AI に与える。

プロンプトは自由

## 汎用性

- 再学習をしなくても、新しいラベルでの物体検出やセグメンテーションが可能

## 効率化

- 訓練データの準備のコストを削減

## 多様な応用

- 自由なラベルに対応できる