

機械学習入門

2019年11月8日 @向陽高校

講師：八谷 大岳

和歌山大学システム工学部 講師
理研革新知能統合研究センター 客員研究員
株式会社サイバーリンクス：顧問

ティーチングアシスタント：黒良 峻平

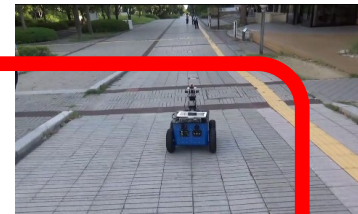
和歌山大学システム工学部4年生

研究テーマ： 実応用指向の機械学習のアルゴリズム研究

ロボット

屋外の移動ロボットの自律ナビゲーション

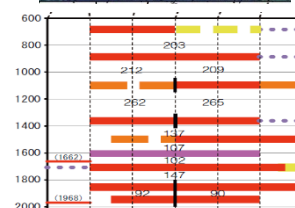
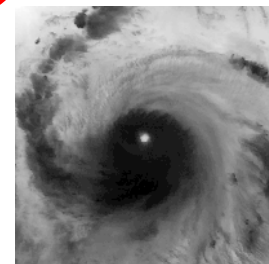
- 株式会社リバストとの共同研究
- ハイブリッドナビゲーション、距離推定、地図圧縮



自然科学

地震・気象の予報モデルの構築

- 理研からの委嘱研究（気象庁、防災研究所）
- 南海トラフモデル同定、地震予報、急発達台風予報、長周期・短周期震動変換など



推薦システム

ファッションコーディネートのおすすめ

- zozoテクノロジーズとの共同研究
- アイテムの集合間マッチング

Query



Answers



つくばチャレンジ

- 遊歩道、公園、ショッピングモール等の実環境のなか、ロボットに自律走行させる技術チャレンジ
 - タスク1: 2kmのコースを完走
 - タスク2: コース内の対象人物（4人）を探索



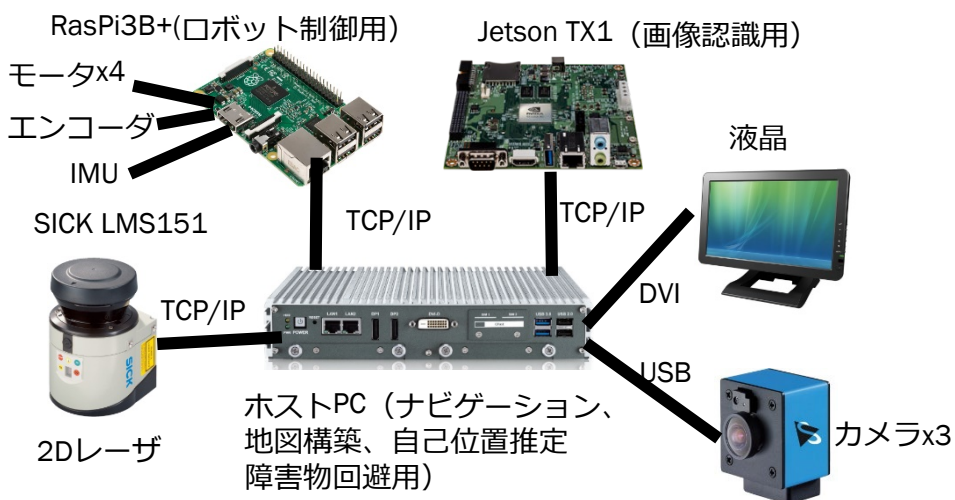
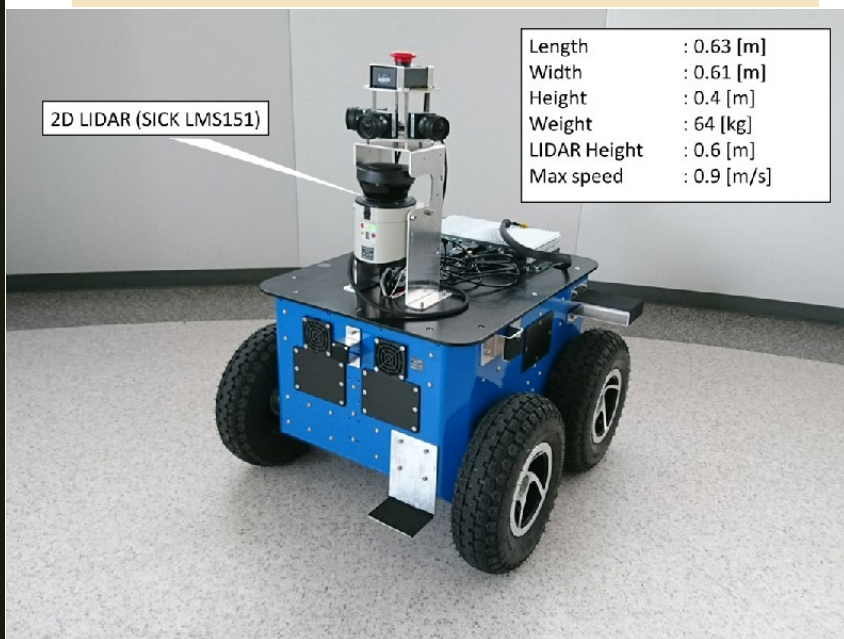
動画 : <https://www.youtube.com/watch?v=SpAwDAaNUy0> <動画>

- 自動運転が注目されるなか、毎年出場ロボットが増加
 - 2015年 : 50台、2016年 : 61台、2017年 : 65台
 - うち完走は9台(2017年)

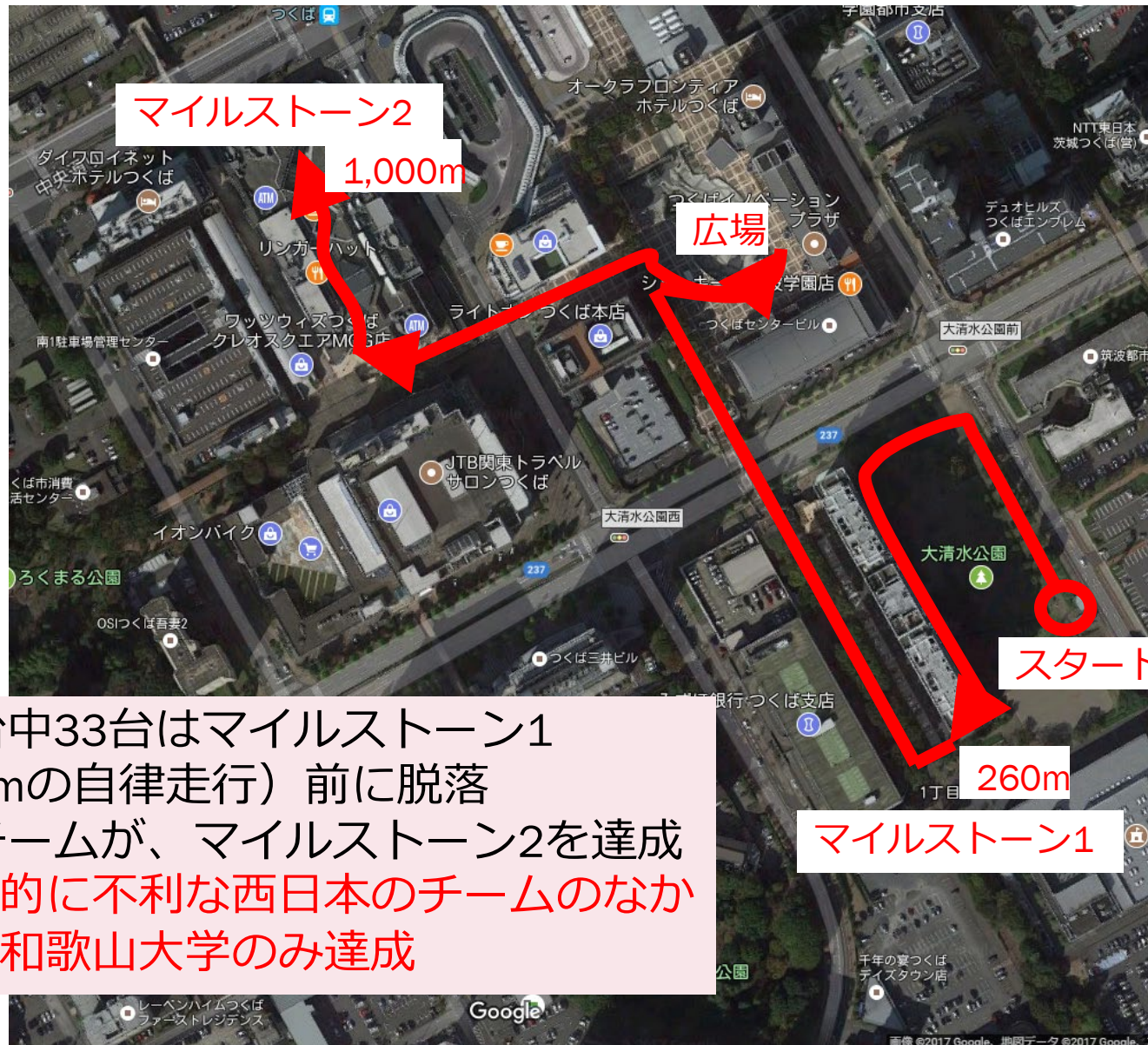
和歌山大学リバストチーム

- 和歌山大学は、株式会社リバストと共同で2017年初出場
- ロボットとシステム構成：
 - Robot OS(ROS)システムを採用。複数のノードをLANで接続し、各ノード（ロボット制御、画像認識、自己位置推定など）を分散処理し、必要に応じて通信

四輪駆動移動型ロボット「Mercury」



マイルストーン2(1,000mの自律走行)を達成！

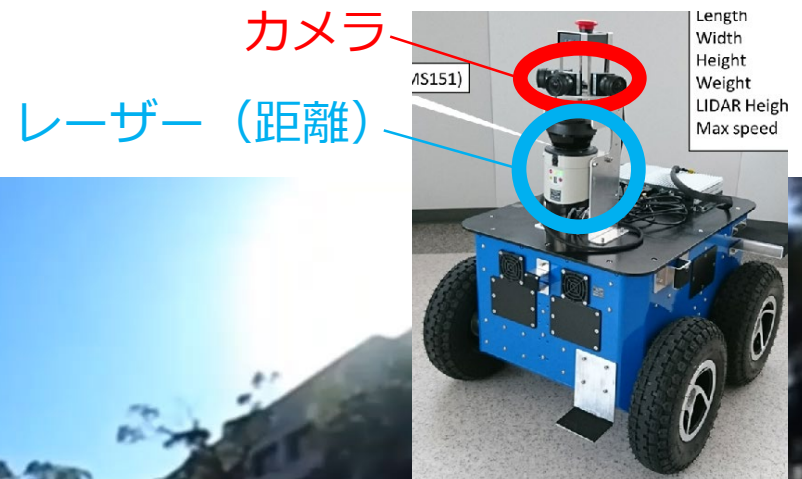


- ・ 65台中33台はマイルストーン1 (260mの自律走行) 前に脱落
- ・ 22チームが、マイルストーン2を達成
- ・ 地理的に不利な西日本のチームのなかでは、和歌山大学のみ達成

つくばチャレンジでの走行の様子

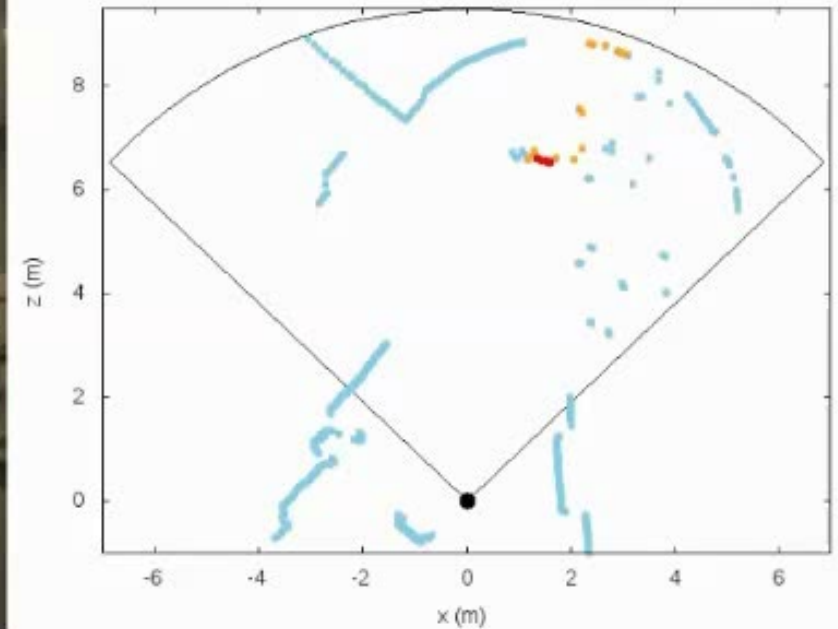
和歌山大学リバストチーム
@つくばチャレンジ2017

ロボットの視界の例



人間と同じような視覚情報や、ロボットならではの距離情報をセンサーを介して取得

ロボットは周囲を認識しながら走行



機械学習を用いてカメラ画像から対象人物を検出し距離を測定

内容

9

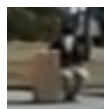
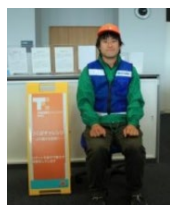
- 機械学習を用いたロボットの自律走行の例
- 機械学習とは
- 機械学習の定式化
- 教師あり学習(予測)の例
- 教師あり学習(分類)の例
- 手書き数字の分類(時間があれば)
- まとめ

機械学習(コンピュータの学習)とは

10

- 人間がコンピュータにルールを与えるのは限界がある

対象人物:

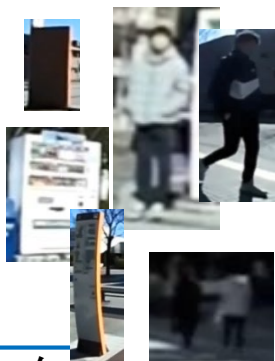
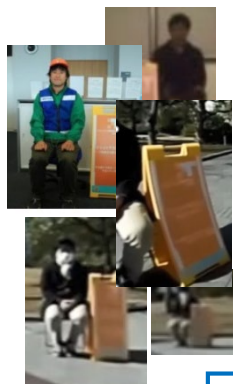


対象人物の見え方は多様!

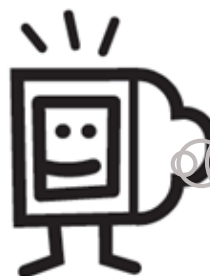
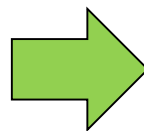
- 機械学習: データから自動でルールを獲得

対象人物の画像

非対象物体の画像



学習データ



機械学習

対象人物

非対象物体

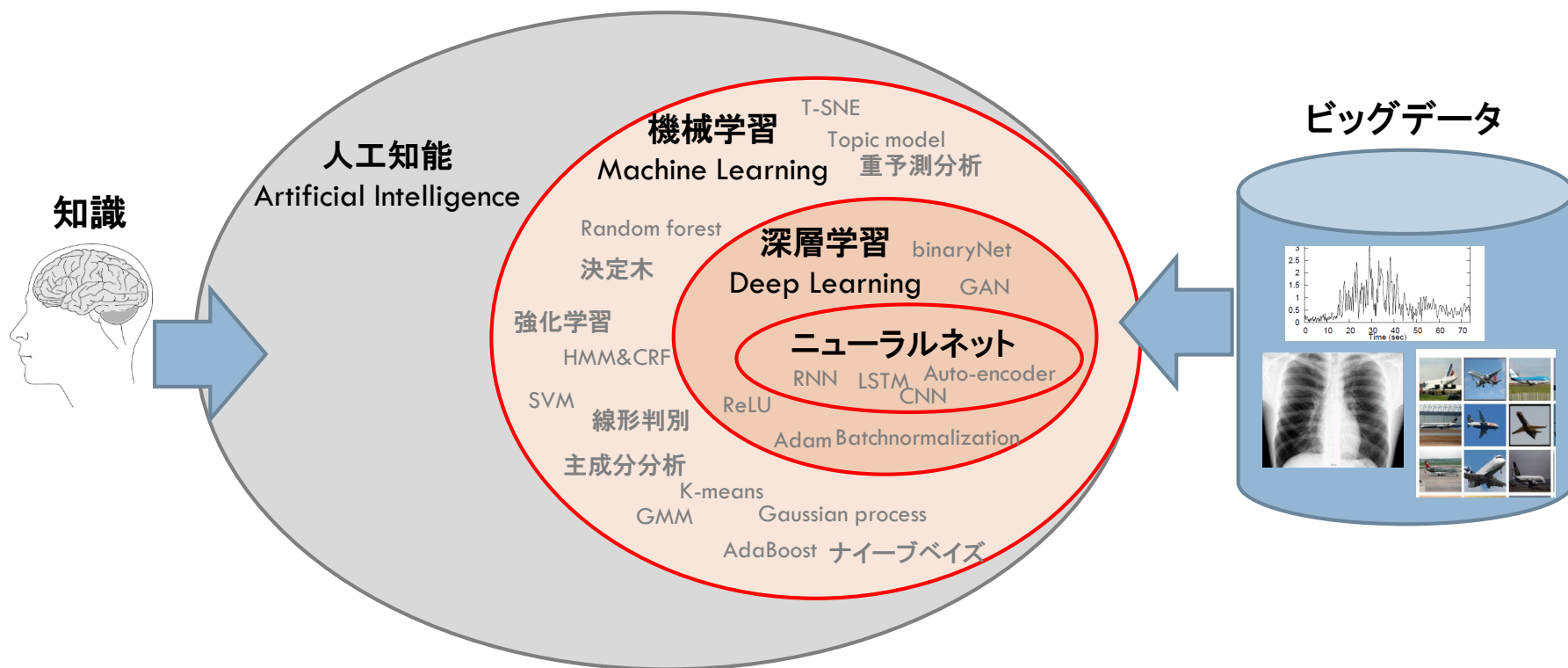
区別するルール
(境界線)

- 2012年頃からビッグデータブームと、深層学習の登場で、様々な分野で一躍注目を集め始める

機械学習と人工知能の関係

11

- **機械学習：人工知能の研究分野の一つ**
 - 人間の知識の活用よりも、データからの知識・ルールの獲得を重視



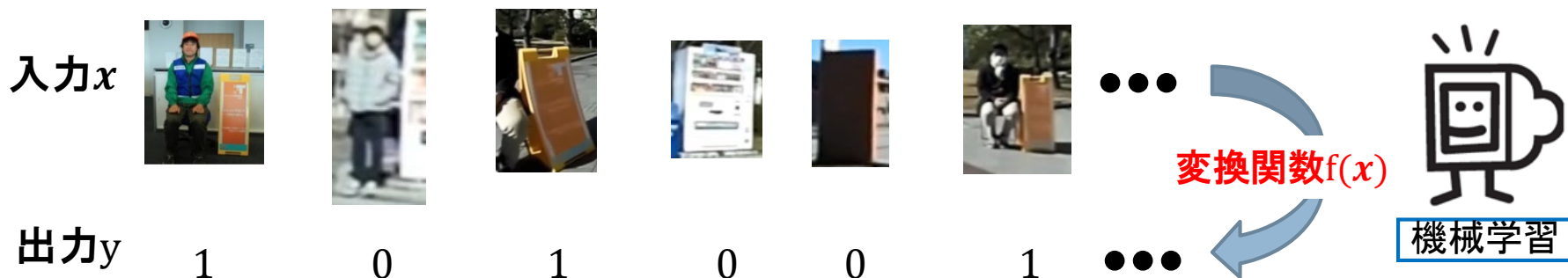
機械学習の種類

12

□ 2種類の学習方法

| 問題 | 定義 | 応用例 |
|--------|---|-----------------------------|
| 教師あり学習 | 入力 x と出力 y の学習データに基づき、入力を出力に変換する関数 $f(x)$ を学習 | 顔認識、株価予測、手書き文字分類、台風の急発達予測など |
| 教師なし学習 | 入力 x のみの学習データに基づき、入力データの傾向を学習 | データの傾向分析 (グループ分け、発生頻度) |

□ 教師あり学習の例



機械学習の種類

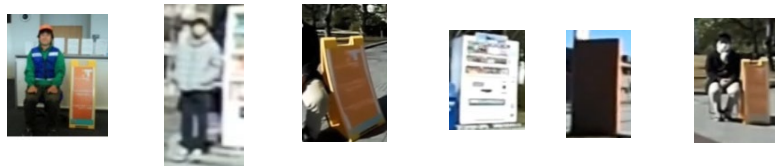
13

□ 3種類の学習方法

| 問題 | 定義 | 応用例 |
|--------|---|-----------------------------|
| 教師あり学習 | 入力 x と出力 y の学習データに基づき、入力を出力に変換する関数 $f(x)$ を学習 | 顔認識、株価予測、手書き文字分類、台風の急発達予測など |
| 教師なし学習 | 入力 x のみの学習データに基づき、入力データの傾向を学習 | データの傾向分析 (グループ分け、発生頻度) |

□ 教師なし学習の例

入力 x



グループ1



グループ2

類似度に基づき
グループ分け



機械学習

演習問題1

14

- 機械学習の教師あり学習と教師なし学習は、
人間や動物の学習方法を参考に構築されました。
- 皆さんが、受験勉強にて実践している過去問を用いた
学習方法で、教師あり学習と教師なし学習に
相当するものがきっとあるはずです。
- 受験勉強における教師あり学習と教師なし学習の例を
10分程度で書いて、メールにて提出してください。

hhachiya@wakayama-u.ac.jp

内容

16

- 機械学習を用いたロボットの自律走行の例
- 機械学習とは
- 機械学習の定式化
- 教師あり学習(予測)の例
- 教師あり学習(分類)の例
- 手書き数字の分類(時間があれば)
- まとめ

教師あり学習：予測

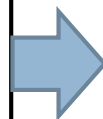
17

- 入力 x と出力 y (**実数値**)のペアの訓練データを準備
 - **学習**: 誤差を最小化するように、数値を予測する関数を学習

【訓練フェーズ】



| CM総時間 x | 販売台数 y |
|-----------|----------|
| 23 | 76 |
| 24 | 77 |
| 29 | 82 |
| 27 | 84 |
| 23 | 74 |
| 24 | 81 |
| 27 | 86 |



入力 x (CM総時間)から
出力 y (販売台数)を予測する
関数を学習

販売台数を予測する関数

$$y \approx f(x)$$

出力: 販売台数 入力: CM総時間

$$\text{予測誤差: } \varepsilon^2 = (y - f(x))^2$$

- **運用**: 学習した関数を用いて、入力に対するラベルを予測

【運用フェーズ】

入力 x
ある週のCM総時間: 30時間



出力 $f(x)$
次の週の販売台数: 90台

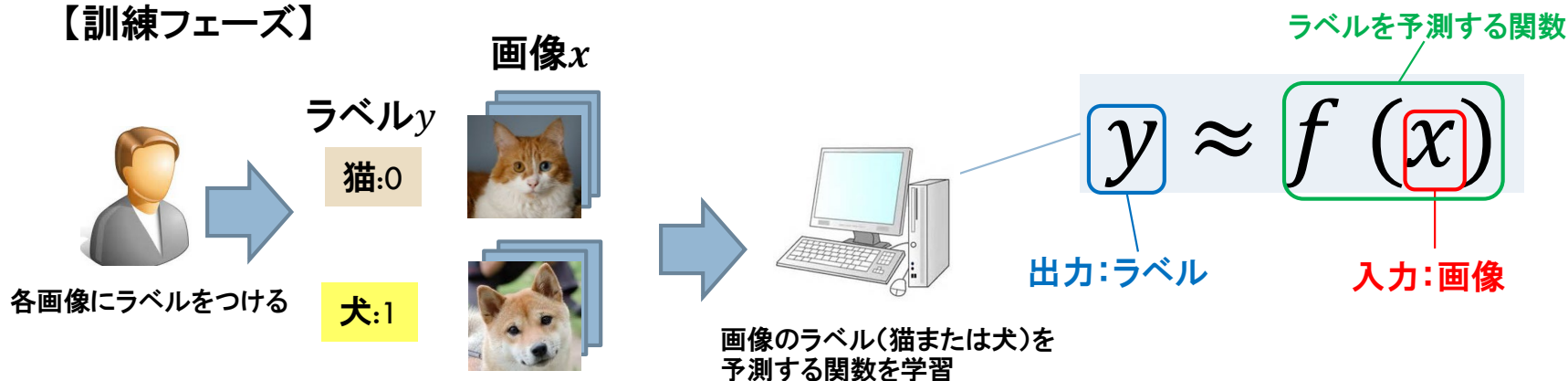
学習した関数を用いて、数値を予測

教師あり学習：分類

18

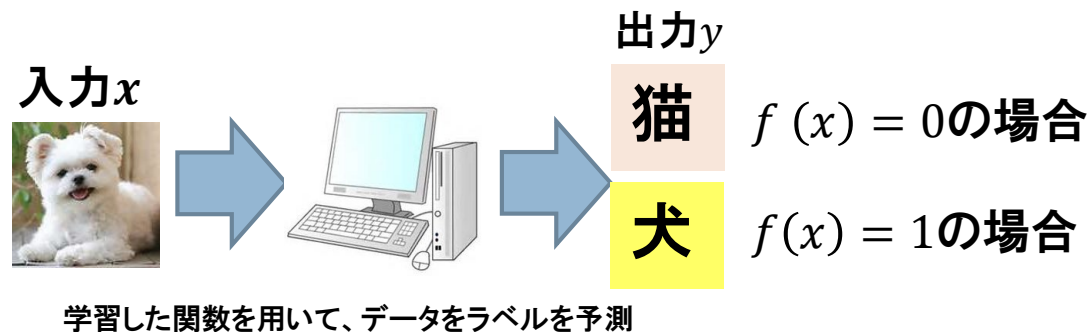
- 入力 x と出力 y (**カテゴリラベル**)のペアの訓練データを準備
 - **学習**: 分類誤差を最小化するように、ラベルを出力する関数を学習

【訓練フェーズ】



- **運用**: 学習した関数を用いて、入力に対するラベルを予測

【運用フェーズ】



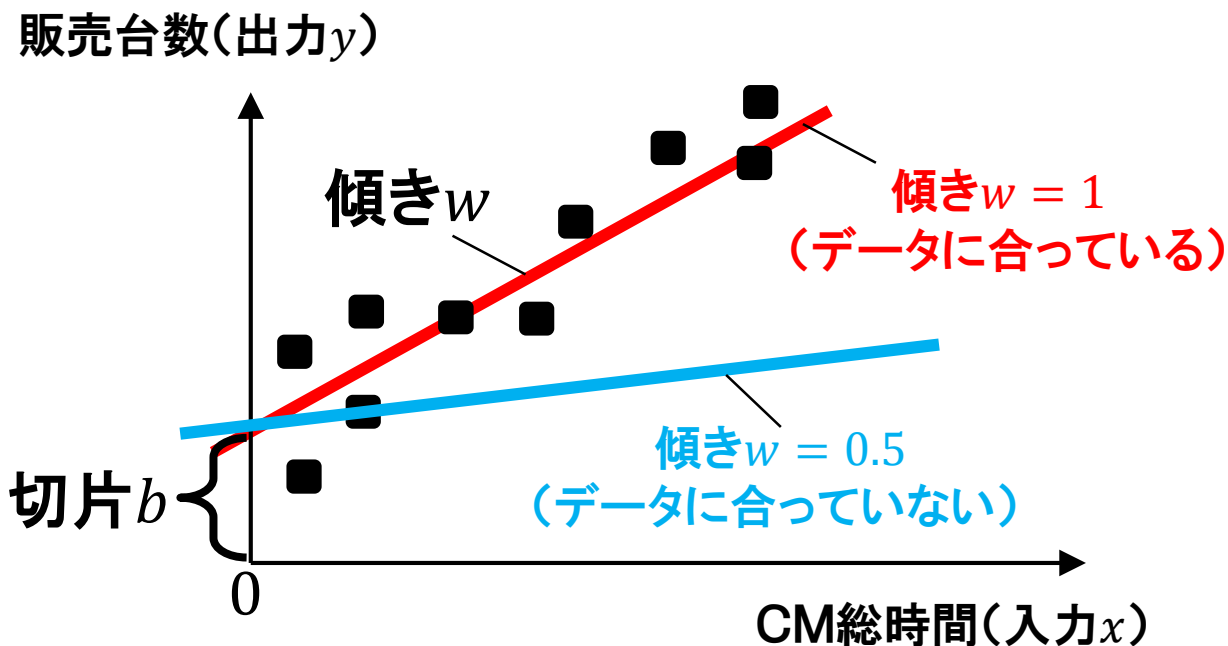
関数 $f(x)$ の基本形

19

□ 予測問題:

$$y \approx f(x) = wx + b$$

直線関数



□ 学習: 傾き w と切片 b をデータに合うように決めること

□ 「データに合う」: 予測誤差 $\varepsilon^2 = (y - f(x))^2$ が小さい

関数 $f(x)$ の基本形 2

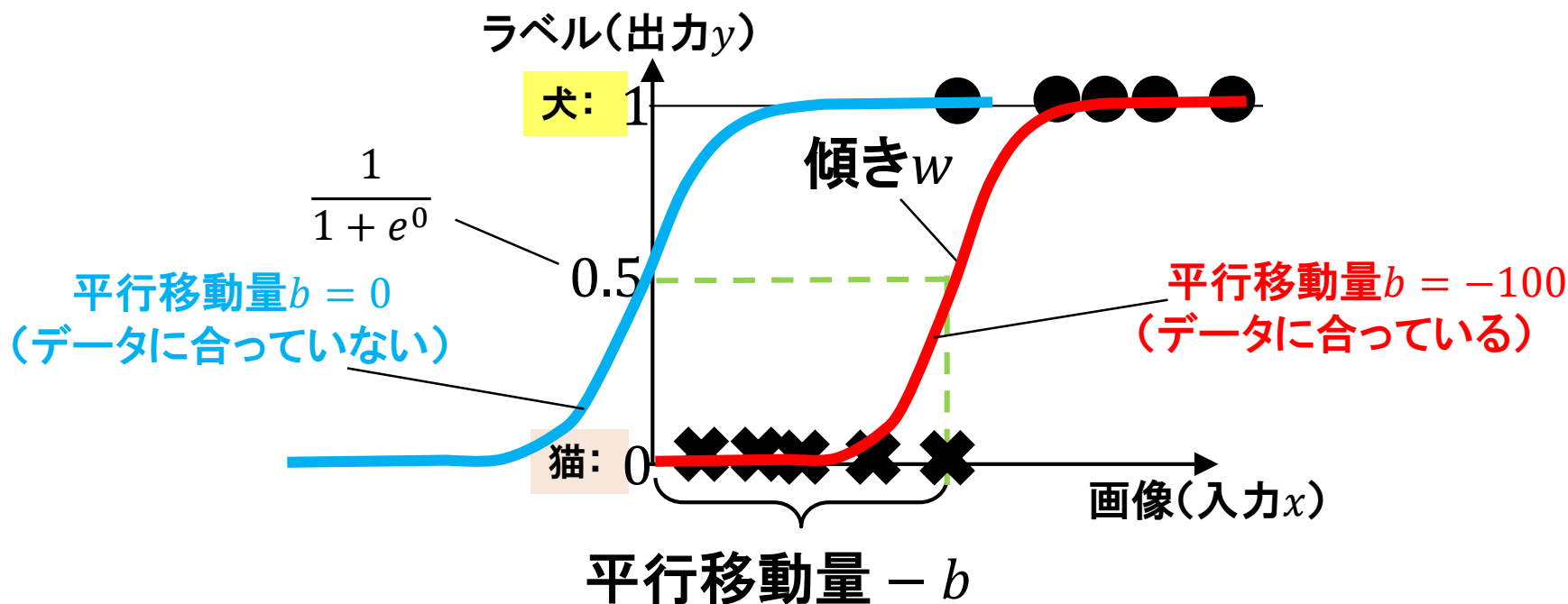
20

□ 分類問題:

$$y \approx f(x) = \frac{1}{1 + e^{-(wx+b)}}$$

シグモイド関数

ネイピア数 $e \approx 2.7$



□ 学習: 傾き w と平行移動量 b をデータに合うように決めること

▣ 「データに合う」: 予測誤差が小さい

内容

21

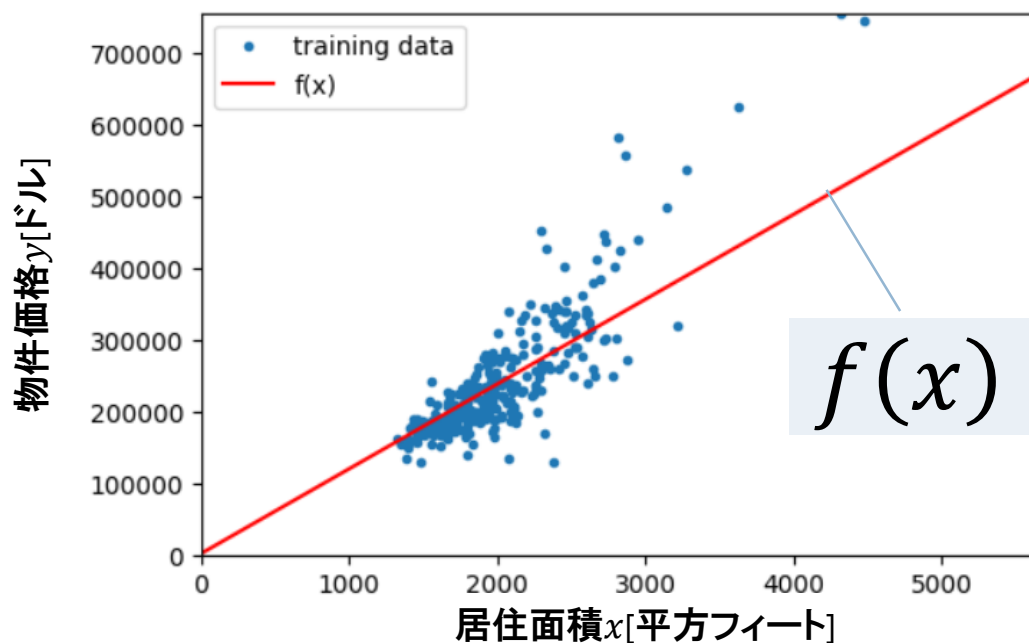
- 機械学習を用いたロボットの自律走行の例
- 機械学習とは
- 機械学習の定式化
- 教師あり学習(予測)の例
- 教師あり学習(分類)の例
- 手書き数字の分類(時間があれば)
- まとめ

演習問題2:ボストンの物件価格

22

- 居住面積(入力 x)から物件価格(出力 y)を予測する
直線関数 $f(x)$ は以下の赤い線のようにになります

1. グラフに基づき傾き w と切片 b のおおよその値を求めましょう
2. 居住面積が3000平方フィートのときの物件価格を予測しましょう



$$f(x) = wx + b$$

- 解答はメールで送ってください

hhachiya@wakayama-u.ac.jp

Pythonによる物件価格予測

23

- 演習問題2の答えは、機械学習に教えてもらいましょう
- 1. 以下のURLをブラウザに入力しアクセスしてください

<https://bit.ly/2NFVQw8>
- 2. ○をクリックして、予測の学習を実行しましょう



Pythonによる物件価格予測 2

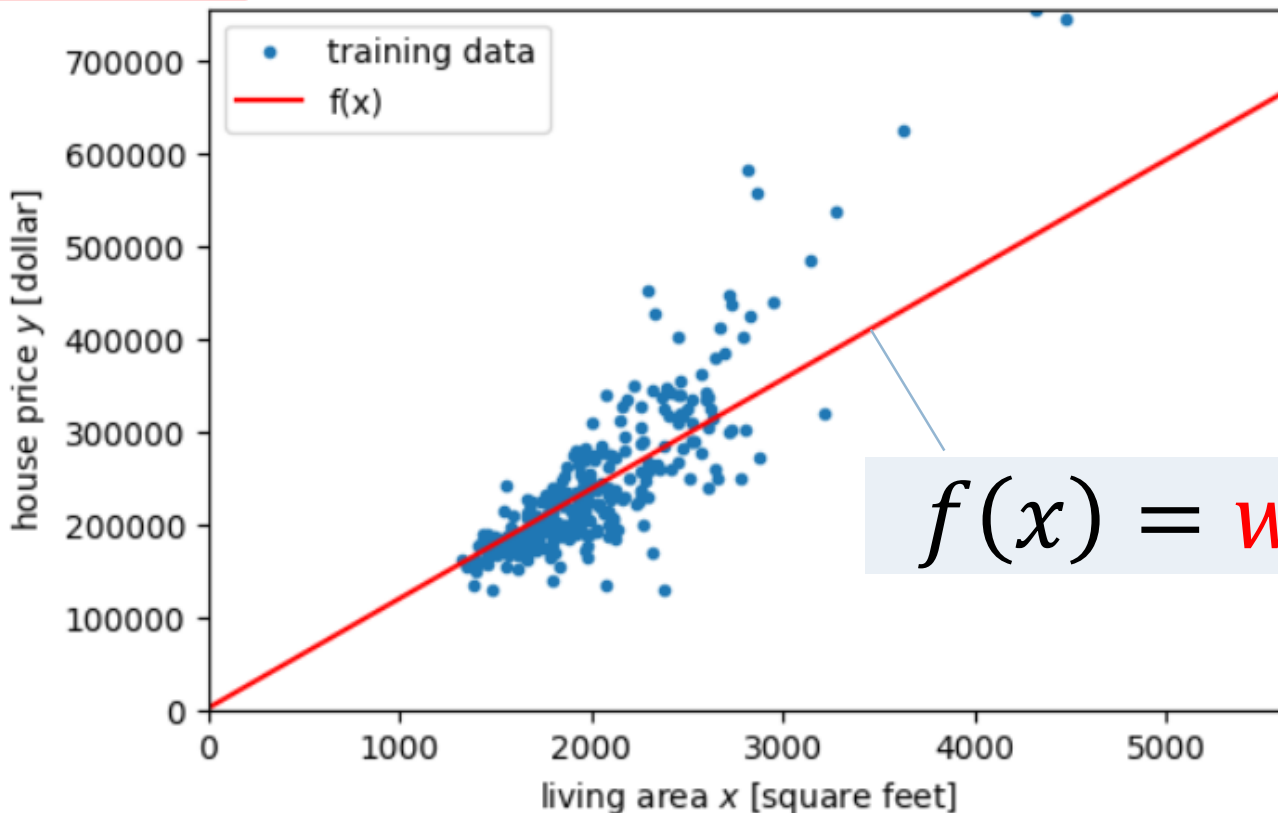
24

3. 学習した傾き w と切片 b を確認しましょう。

学習結果：

傾き w =

切片 b =



$$f(x) = wx + b$$

Pythonによる物件価格予測 3

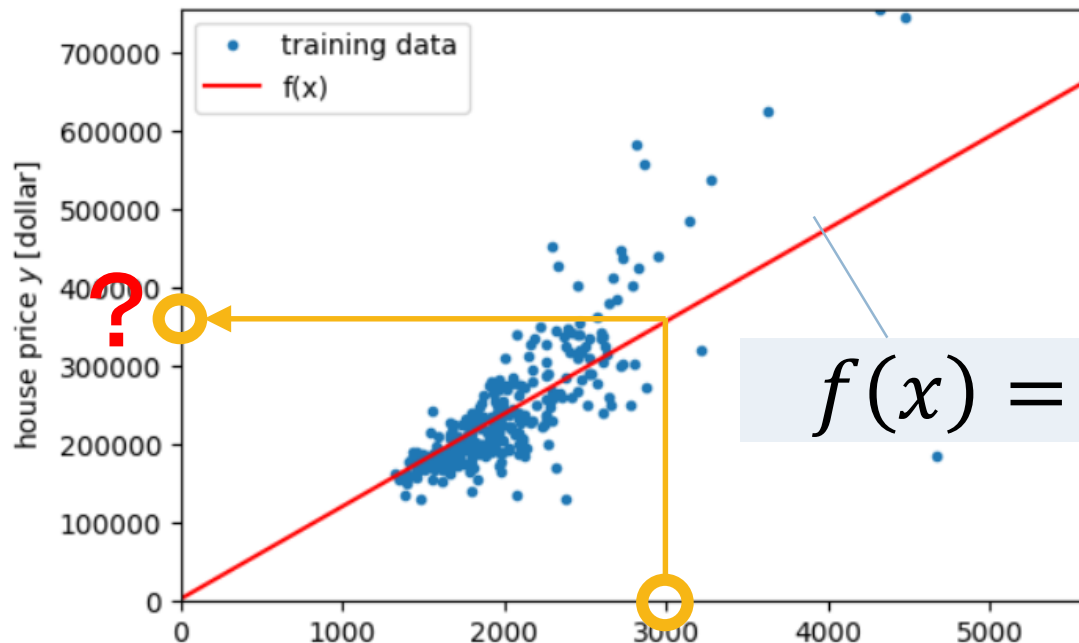
25

4. 居住面積が3000平方フィートのときの物件価格を予測しましょう

- 以下のように「`myML.f(x)`」関数を実行してみましょう



- 機械学習は、学習した関数 $f(x)$ を使って予測してくれます



$$f(x) = wx + b$$

休憩 10分
(〇〇時〇〇分まで)

内容

27

- 機械学習を用いたロボットの自律走行の例
- 機械学習とは
- 機械学習の定式化
- 教師あり学習(予測)の例
- 教師あり学習(分類)の例
- 手書き数字の分類(時間があれば)
- まとめ

関数 $f(x)$ の基本形

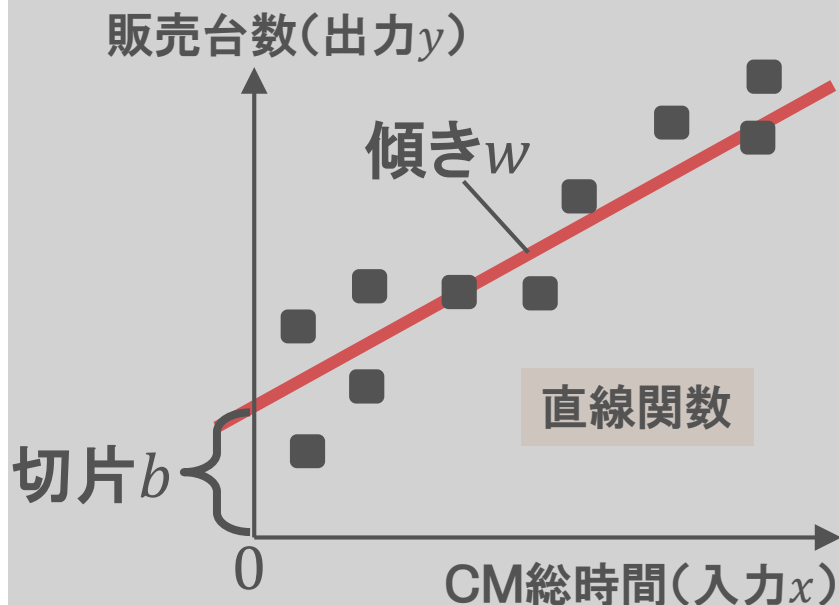
28

□ 予測:

$$y \approx f(x) = wx + b$$

□ 学習対象:

w : 傾き、 b : y の切片



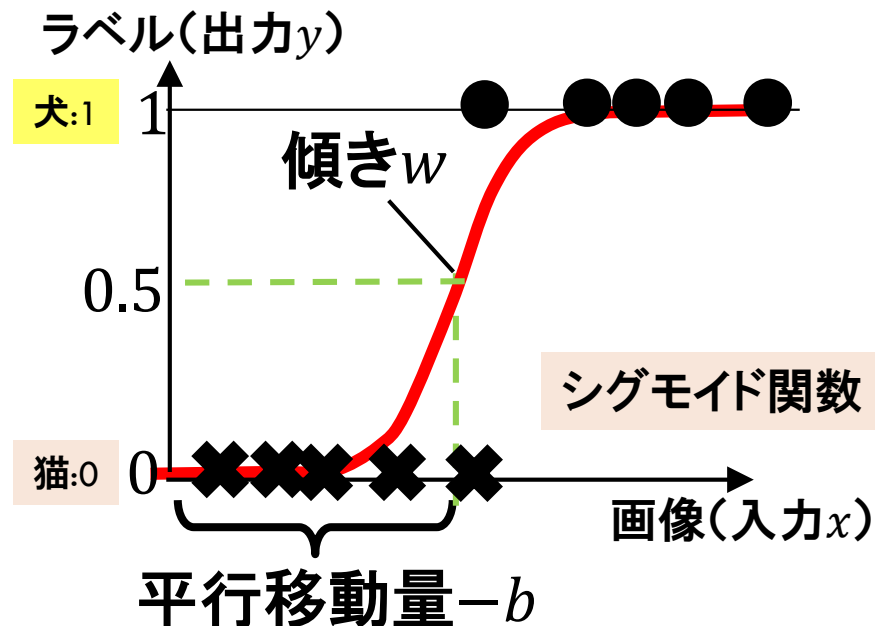
□ 分類:

ネイピア数 $e \approx 2.7$

$$y \approx f(x) = \frac{1}{1 + e^{-(wx+b)}}$$

□ 学習対象:

w : 傾き、 b : 平行移動量

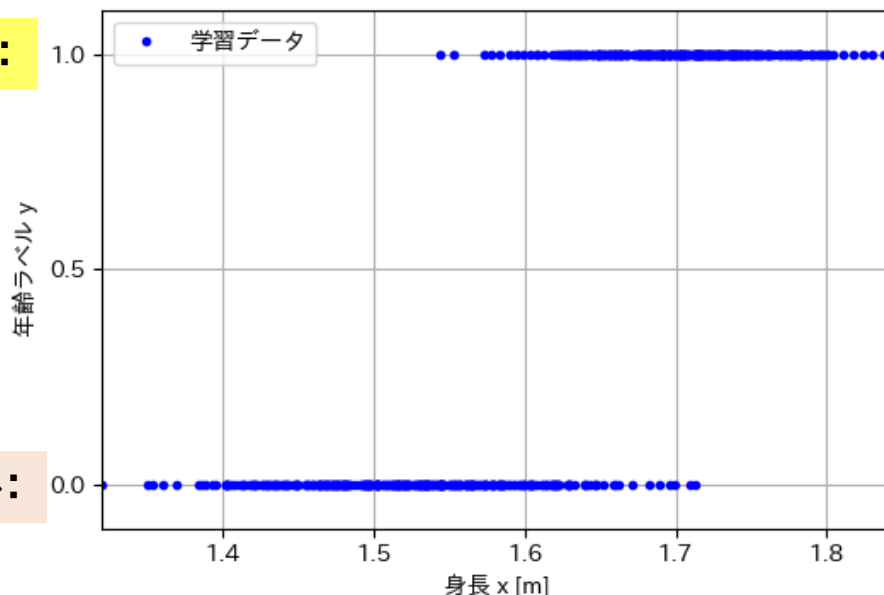


演習問題3: 身長に基づく年齢分類

29

- 身長(入力 x)から年齢ラベル(0:13歳、1:17歳)を分類するシグモイド関数をコンピュータに学習させましょう
 1. 傾き w と平行移動量 b の値を確認しましょう
 2. コンピュータが選んだ17歳と13歳の境界の身長を確認しましょう

17歳のラベル:



13歳のラベル:

PythonコードのURL:
<https://bit.ly/2pAU9br>

解答はメールで送ってください:

hhachiya@wakayama-u.ac.jp

Pythonによる年齢分類

30

1. 以下のURLをブラウザに入力しアクセスしてください

<https://bit.ly/2pAU9br>

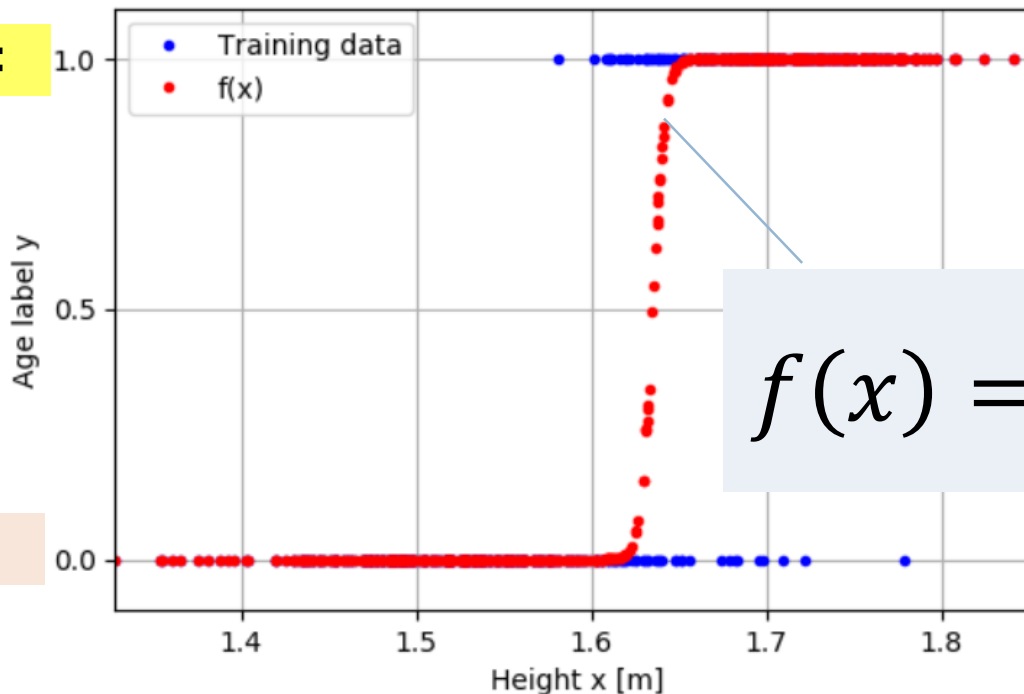
2. 分類の学習を実行し、学習した関数を確認しましょう

学習結果：

傾き w =

切片 b =

17歳のラベル：



$$f(x) = \frac{1}{1 + e^{-(wx+b)}}$$

13歳のラベル：

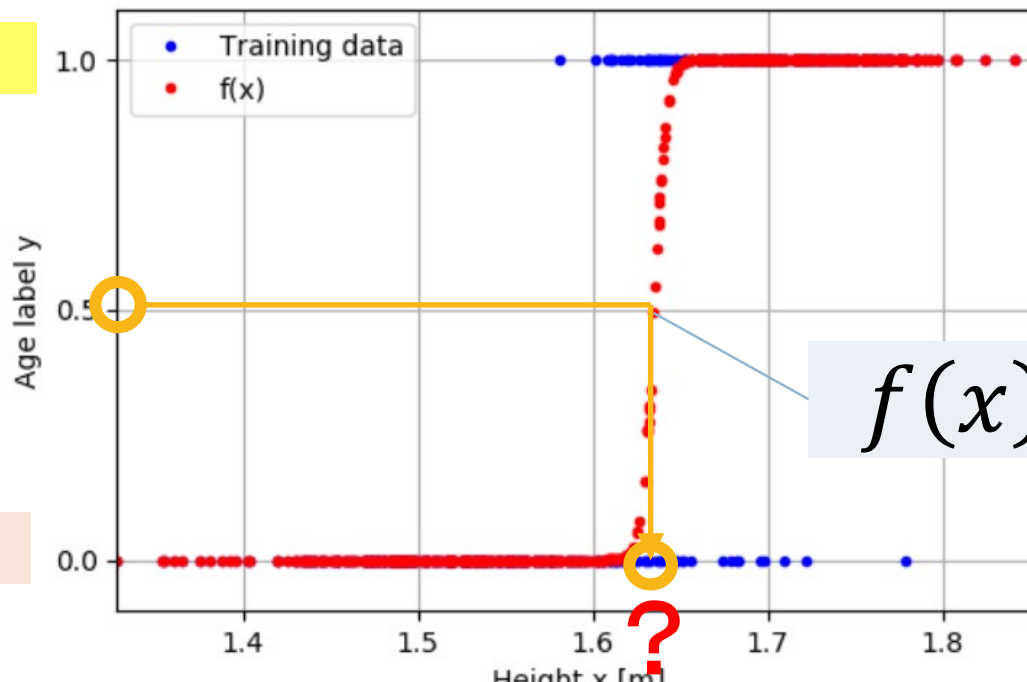
Pythonによる年齢分類 2

31

3. 以下のように「`myML.f(x)`」関数を実行し、出力が0.5近くになるところを探してみましょう。

▶ `myML.f(1.6)`

17歳のラベル:



13歳のラベル:

$$f(x) = 0.5$$

分類の応用例1

32

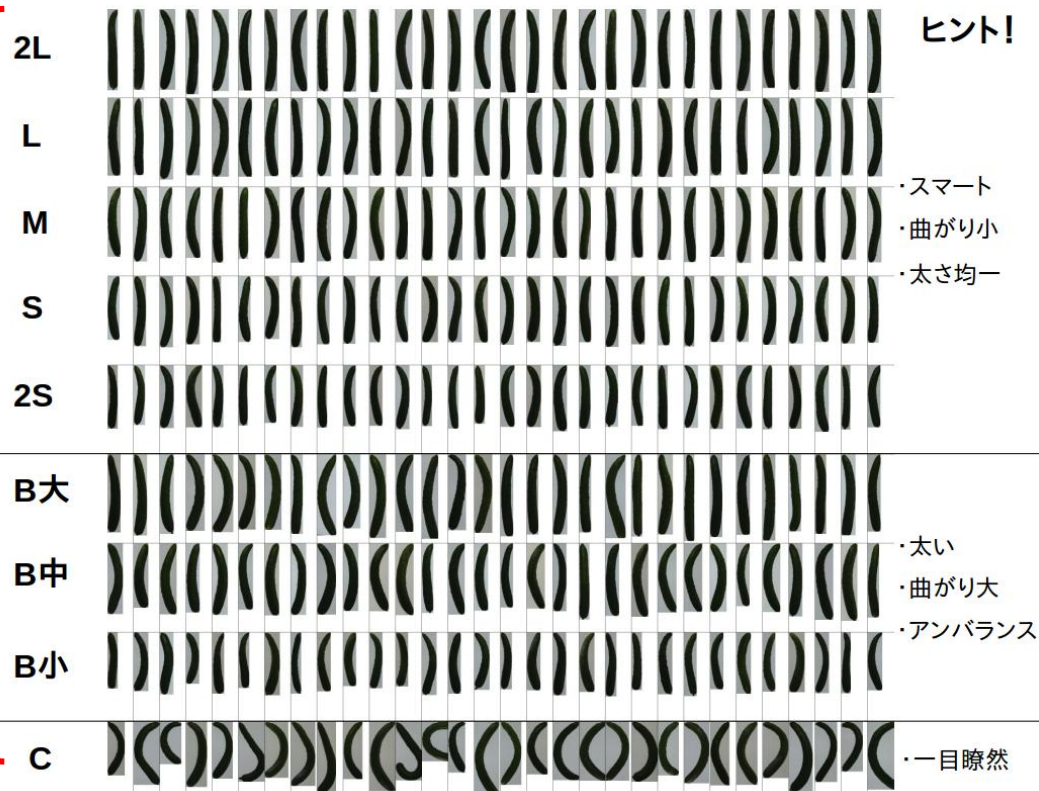


<https://www.youtube.com/watch?v=XkKxSAb4EAw>

きゅうりの等級分類

33

9種類の仕分けカテゴリ



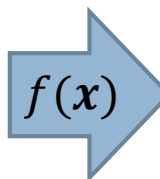
大きさ、形、太さ
の組み合わせに
よって決まる

<http://workpiles.com/2017/08/ccb9-proto3-description2/>

入力 x : 画像



...



C L S B小 ...

出力 y : 等級 $\in \{2L, L, M, \dots, C\}$

分類の応用例2

34

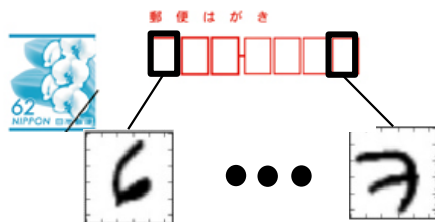
□ 郵便番号認識

- 1時間に4万通の郵便番号を認識し、地域ごとの仕分けが可能

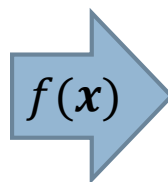


<https://jpn.nec.com/kids/himitsu/04.html>

□ 手書き数字の分類



入力 x : 各文字部分の画像



$\boxed{6} \dots \boxed{7}$

出力 y : 数字 $\in \{0, 1, \dots, 9\}$

内容

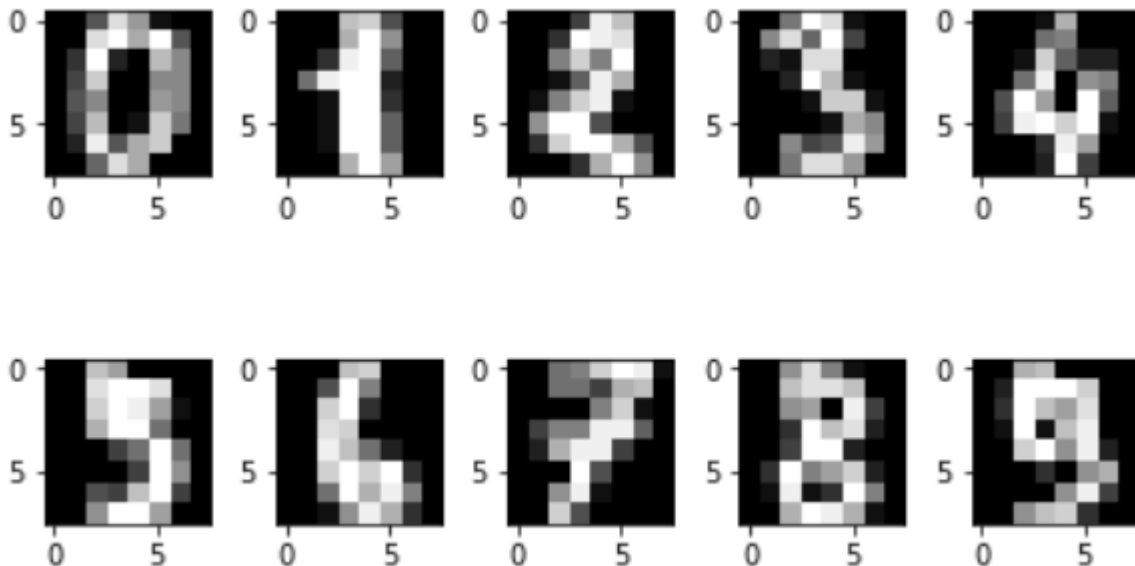
35

- 機械学習を用いたロボットの自律走行の例
- 機械学習とは
- 機械学習の定式化
- 教師あり学習(予測)の例
- 教師あり学習(分類)の例
- 手書き数字の分類(時間があれば)
- まとめ

Pythonによる手書き数字分類

36

- 以下のような0～9の手書きの数字を、コンピュータに分類させてみましょう



- 以下のURLをブラウザに入力しアクセスしてください

<https://bit.ly/2oXfMIS>

Pythonによる手書き数字分類 2

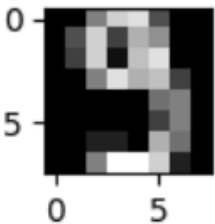
37

- デフォルトでは、数字の「7」と「9」を分類するように設定
- 実行すると、学習後に、最大5個まで成功例と失敗例が表示

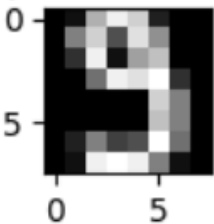
-----成功例-----

72個中70個が分類成功

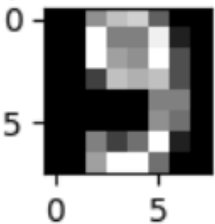
pre:9, true:9



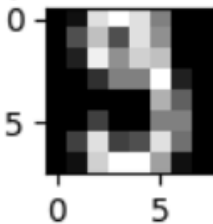
pre:9, true:9



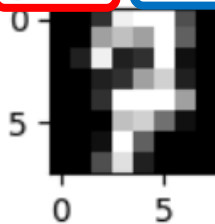
pre:9, true:9



pre:9, true:9



pre:7, true:7



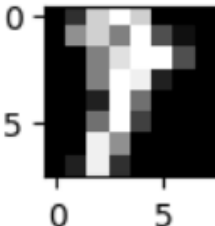
予測した数字

本当の数字

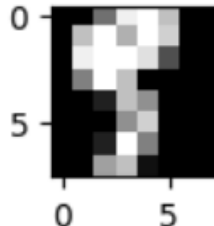
-----失敗例-----

72個中2個が分類失敗

pre:9, true:7



pre:7, true:9



成功または
失敗した個数

Pythonによる手書き数字分類 2

38

- 分類対象の数字を変えて、どの組み合わせが一番難しいのか調べてみましょう

```
#####  
# メイン  
  
#-----  
# データの準備  
label0 = 7  
label1 = 9  
  
# データの読み込み  
MNIST = datasets.load_digits()  
X = MNIST.data[(MNIST.target==label0) | (MNIST.target==label1)]  
Y = MNIST.target[(MNIST.target==label0) | (MNIST.target==label1)][np.newaxis].T  
Y[Y==label0] = 0  
Y[Y==label1] = 1  
  
trNum = int(len(X)*0.8)  
Xtr = X[:trNum,:]  
Ytr = Y[:trNum,:]  
Xte = X[trNum:,:]  
Yte = Y[trNum:,:]  
#-----
```

0～9までの異なる数字を設定してみましょう

解答はメールで送ってください:

hhachiya@wakayama-u.ac.jp

まとめ

39

- 機械学習とは
 - 人間や動物の学習方法を参考にしてコンピュータの学習方法を構築
 - 機械学習の種類: 教師あり、教師なし
- 機械学習の定式化
 - 学習データから関数を学習する問題として定式化
 - 予測と分類などタスクに合わせて関数(直線、シグモイド)を定義
- 教師あり学習(予測)&教師あり学習(分類)の例
 - 学習データに合うように関数のパラメータを手動およびpythonで調整
 - 農業やはがきの分類などの応用例を紹介
- 手書き数字の分類(時間があれば)
 - 実際に手書き分類がどれくらいできるのかを確認