

実データで学ぶ人工知能講座
講座準備（初回）

マシュー ホーランド
Matthew J. Holland
`matthew-h@ar.sanken.osaka-u.ac.jp`

大阪大学 産業科学研究所 助教

講座資料のネット上の居場所

演習課題のソフトウェア

Python のソースファイルと Jupyter ノートブックを下記から入手.

<https://feedbackward.com/content/cds/practice.tar.gz>

UNIX 系では, `tar -xzf <FILE>` で解凍.

自己紹介

Matthew Holland
大阪大学 産業科学研究所 助教

出身地
Calgary (カナダ)

略歴

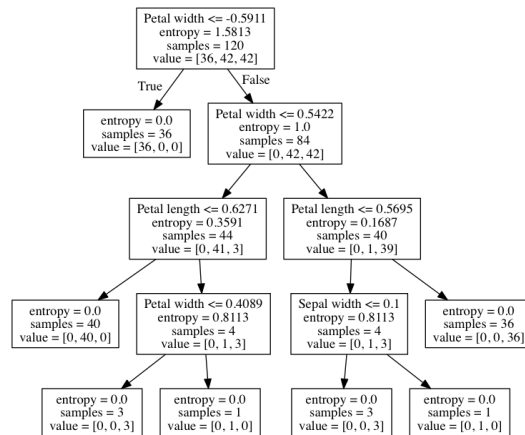
- ▶ NAIST 数理情報学研究室 (2013/04–2017/12)
- ▶ JSPS 特別研究員 (2015–2017)
- ▶ 大阪大学 IDS (2018/01–2019/03)
- ▶ 大阪大学 産研 (現職) (2019/04–...)

自己紹介

ここ数年の研究課題など

- ▶ Feedback, computation, and robust learning algorithms (博論)
- ▶ ロス関数の自動選択アルゴリズムと系列データへの拡張の解析 (JSPS 特別研究員)
- ▶ 安全な AIこそ効率的: ロバスト学習による汎化性能向上の研究 (JST ACT-I)
- ▶ 識別マージンの強い分布的統制による頑健化と効率化 (科研費, AISTATS 2019)
- ▶ ロバスト勾配降下法 (ICML 2019, AISTATS 2019)
- ▶ PAC-Bayes の効率的な頑強化 (NeurIPS 2019)
- ▶ 最近: ロバストな確率的勾配降下法や非リスク (CVaR 等) 学習則など...

講座の概要（マシュー担当分）



講座の概要（マシュー担当分）

```
# エントロピーを指標とする決定木のインスタンスを生成し、
# 決定木のモデルに学習データを適合させる
tree = DecisionTreeClassifier(criterion='entropy', max_depth=3, r
tree.fit(X_train_std, y_train)
```

講座の概要（マシュー担当分）

```
# テストデータ分割のための乱数のシード（整数値）
random_seed = 2
# テストデータの割合
test_proportion = 0.2
# Iris データセットをロード
iris = datasets.load_iris()
# 特徴ベクトルを取得
X = iris.data
# クラスラベルを取得
y = iris.target

# 学習データとテストデータに分割
X_train, X_test, y_train, y_test = train_test_
irition, random_state = random_seed)
```

講座の概要（マシュー担当分）

```
#### 決定木学習による識別と決定木の描画
#### Python機械学習本：3.6.2節（pp. 84-86）

from sklearn import datasets
import numpy as np
from matplotlib.colors import ListedColormap
import matplotlib.pyplot as plt
# for scikit-learn <=0.17
#from sklearn.cross_validation import train_test_split
# for scikit-learn 0.18
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.tree import DecisionTreeClassifier, export_graphviz
from sklearn.metrics import precision_recall_fscore_support
```

講座の概要（マシュー担当分）

本講座の方針：

“自分で実装してはじめて，機械学習の本質が見えてくる”

※ 但し，一定の基礎知識を持つことが大前提。

実施日程

7/11 (土)，7/25 (土)，8/1 (土)。

講座の概要（マシュー担当分）

Day 1「土台作り」

▶ 学習とは何か？

付属演習：PLA の解剖

▶ ノイズとフィードバックと汎化能力

付属演習：ノイズのある線型識別

▶ 長方形の例から AI の性能保証を考える

付属演習：長方形の学習

▶ 表現力と性能保証

付属演習 (1)：Hoeffding の不等式の成否を検討，(2)：PLA の性能保証：理論と現実

【中入り】（視覚情報エンコーダ等の AI 応用例）

講座の概要（マシュー担当分）

Day 2「データから情報へ」

▶ フィルターバンクと特徴量づくり

付属演習：Gabor フィルターバンクによる特徴量づくり

▶ データ紹介 (vim-2)：視覚情報エンコーダ学習用のデータセット

Day 3「手法の理解を本物に」

▶ 凸性と最適化と学習

付属演習：最適化アルゴリズムの数値検証

▶ 線型モデルの学習（実数値の予測）

付属演習：Legendre 多項式を用いた線型回帰

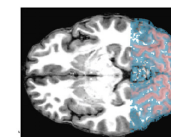
▶ 総合演習：視覚情報エンコーダの学習

脳の符号化と復号化 (Nishimoto et al., 2011)

自然動画 ($20^\circ \times 20^\circ$ @ 15Hz)



脳活動 (fMRI による BOLD 信号, 1Hz)



参考文献

Nishimoto, S., Vu, A. T., Naselaris, T., Benjamini, Y., Yu, B., and Gallant, J. L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, 21(19):1641–1646.