

# 授業準備：Webclassからコードをダウンロードし、 Google colaboryで開いておいてください

## 演習授業中の質問対応について

The screenshot shows a Zoom meeting window. The main video area displays a slide with the text: "演習授業中の質問をチューターの先生が対応させていただきます。" (We will respond to questions during the practice lesson by the tutor teacher.) Below this, the name "曹日昌" is visible. A red box highlights the reaction bar at the bottom of the video, which includes icons for thumbs up, thumbs down, and a hand icon. A red arrow points from the text "演習にエラーが出たなど問題があったらリアクションの挙手を押してください。" (If there is an error during the practice, please press the reaction hand icon.) to the hand icon. Another red box highlights the chat window on the right, which has the title "ミーティング チャット". A red arrow points from the text "質問内容を入力して、「全員」宛てに送信してください。" (Enter the question content and send it to "Everyone".) to the "宛先" (To) dropdown menu in the chat, which is currently set to "全員" (Everyone).

Zoom ミーティング

演習授業中の質問をチューターの先生が対応させていただきます。

曹日昌

ミーティング チャット

演習にエラーが出たなど問題があったらリアクションの挙手を押してください。

質問内容を入力して、「全員」宛てに送信してください。

Miho Ishimaru

ミュート解除 ビデオの開始 セキュリティ 参加者 画面共有 リアクション アプリ ホワイトボード ノート 詳細 終了

メッセージは誰に表示されますか？  
宛先 全員

ここにメッセージを入力します...

# 医療とAI・ビッグデータ入門

## 演習20

### まとめの演習

# 演習20 課題

Webclassで課題を提出してください。締め切りは**2024/02/15 23:59**まで

**(必修課題)** Heart Attack dataset を使い、分類を行なってください

- 使用するモデルは演習10-演習14で取り扱った、LogisticRegression, SVC, DecisionTreeClassifier, RandomForestClassifierのうち、どれか一つを選択してください。
- webclassからkadai\_ML.ipynbをダウンロードし、空欄となっているコード5~コード8にコードを書き、実行できることを確認してください。
- kadai\_ML\_(学生番号).ipynbに名前を変更し、webclassで提出してください。

**(発展課題)** 全員が提出する必要はありません。提出された場合は成績に加点します。

皮膚がんの画像データを使って、分類を行なってください。

- ニューラルネットワークで自分でモデルを設計してください。
- webclassからkadai\_DL.ipynbをダウンロードし、空欄となっているコード6~コード8にコードを書き、実行できることを確認してください。
- kadai\_DL\_(学生番号).ipynbをwebclassで提出してください。

# 演習20 課題

Heart Attack dataset を使い、分類を行なってください

コード3

heart.csvをGoogle Driveにアップロードして、pd.read\_csv()で読み込む

```
df = pd.read_csv('/content/drive/MyDrive/heart.csv')
df
```

\* Google Driveの MyDrive直下にheart.csvをおいている必要があることに注意

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

# 演習20 課題

Heart Attack dataset を使い、分類を行なってください

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2	0

303 rows × 14 columns ➡ 303行 × 14列 (13個の特徴量+1個の正解値)

## 演習20 課題

**Heart Attack dataset** を使い、分類を行なってください

age sex cp trtbps chol fbs restecg thalachh exng oldpeak slp caa thall output

**age** : 年齢

**sex** : 性別

**cp** : Chest pain type (0=典型的狭心症、1=非定型狭心症、2=非狭心症性疼痛、3=無症状)

**trtbps** : 安静時血圧

**chol** : コレステロール値

**fbs** : 空腹時血糖 120mmHg/dl以上の時1、以下の時0

**restecg** : 安静時心電図結果~0 = 正常、1 = ST-T波正常、2 = 左室肥大

**thalachh** : 最大心拍数

**exng** : 運動誘発狭心症~1 = あり、0 = なし

**oldpeak** : 安静時に比べて運動により誘発されるST低下

**slp** : 運動ピークSTセグメントの傾き : 0 : 上り勾配、1 : 平坦、2 : 下り勾配

**caa** : 主要血管の数 (0-3)

**thall** : タリウム負荷試験 : 0 : 正常0、1 : 正常1、2 : 固定欠損、3 : 可逆欠損

**output** : ターゲット変数 (1=心臓発作の可能性が高い)

# 演習20 課題

Heart Attack dataset を使い、分類を行なってください

コード4 y\_data (正解値データ)、x\_data (特徴量データ)を作成する

```
y_data = df.iloc[:, 13]
x_data = df.iloc[:, 0:13]
```

y\_data

```
0    1
1    1
2    1
3    1
4    1
```

```
..
298  0
299  0
300  0
301  0
302  0
```

Name: output, Length: 303, dtype: int64

y\_data.shape

(303,)

x\_data

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	3
299	45	1	3	110	264	0	1	132	0	1.2	1	0	3
300	68	1	0	144	193	1	1	141	0	3.4	1	2	3
301	57	1	0	130	131	0	1	115	1	1.2	1	1	3
302	57	0	1	130	236	0	0	174	0	0.0	1	1	2

303 rows x 13 columns

x\_data.shape

(303, 13)

## 演習20 課題

Heart Attack dataset を使い、分類を行なってください

kadai\_ML.ipynbのコード5~8 に以下のコードを書いて実行してください。

コード5 : `train_test_split()`で学習データと検証データに分割してください

コード6 : 学習モデルを選択してください(`LogisticRegression`, `SVC`,  
`DecisionTreeClassifier`, `RandomForestClassifier`のうちどれか一つ)

\*`LogisticRegression`で行う場合には、収束しないエラーが起こることがあり、その場合は、  
(モデル名) = `LogisticRegression(max_iter=1000)`でインスタンス化を行なってください。

コード7 : 学習データを入れて、学習を行なってください

コード8 : モデルの評価を行なってください(正解率、AUC、precisionなど好きなものの一つ以上)

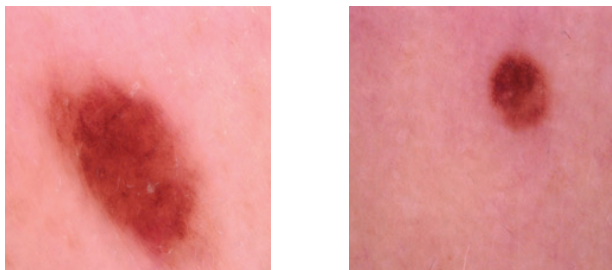


## 演習20 課題

**(発展課題) 皮膚がんの画像データを使って、分類を行なってください。**

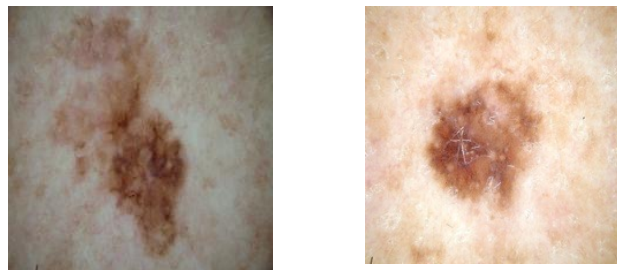
1) images\_skin\_cancer.zipをGoogle Driveにアップロードしてください。

良性腫瘍(正解値 : 0)



**1440枚のデータ**

悪性腫瘍 (正解値 : 1)



**1197枚のデータ**

出典(<https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>)

## 演習20 課題

(発展課題) 皮膚がんの画像データを使って、分類を行なってください。

コード1 : Google Driveにマウントする

```
from google.colab import drive  
drive.mount('/content/drive', force_remount=True)
```

コード2 : images\_skin\_cancer.zip を解凍する

```
!unzip '/content/drive/MyDrive/images_skin_cancer.zip'
```

\* Google Driveの MyDrive直下にimages\_skin\_cancer.zipをおいている必要があることに注意

(発展課題) 皮膚がんの画像データを使って、分類を行なってください。

コード4：前処理 (このまま全部実行してください)

- 画像データを読み込み、x\_trainとy\_trainを作成する
- 今回はカラーモード、サイズは64×64で読み込む

## 演習20 課題

(発展課題) 皮膚がんの画像データを使って、分類を行なってください。

コード5 ランダムシード値の設定

```
set_random_seed(0)
```

## 演習20 課題

**(発展課題) 皮膚がんの画像データを使って、分類を行なってください。**

**kadai\_DL.ipynbのコード6~8に以下のコードを書いて実行してください。**

**コード6 モデルを設計してください**

(中間層は2層以上でニューラルネットワークで作成してください)

**コード7 学習をしてください(モデル名.fit())**

(batch\_size =, epochs =, validation\_split =の値を自分で決める)

\*かなり時間がかかるので、epochs数は10~20ぐらいから始めるのをお勧め

**コード8 accuracyとval\_accuracyの学習過程の図示、lossとval\_lossの学習過程の図示**

演習を始める前に2つアナウンス

医療とAI・ビッグデータ応用のご案内

# 「医療とAI・ビッグデータ応用」の履修について

■開講時期:前期 2024年度は4月25日～7月18日の予定

■単位数:1単位

■対象学科

必修科目 :医学科2年、歯学科2年

選択必修科目:保健衛生学科・検査技術学専攻2年～

自由科目 :口腔保健学科・口腔保健衛生学専攻/口腔保健工学専攻 2年～  
保健衛生学科・看護学専攻 2年～

■履修方法:

必修科目以外については、4月初旬に各教務係から案内があります。

希望者は申し込みをしてください。

演習を始める前に2つアナウンス

授業評価アンケート



# 授業評価アンケートについて



Webclass QRコード

学科・専攻	コース番号
医学科	IL2300828
看護学専攻	IL2300325
検査技術学専攻	IL2300332
歯学科	IL2300143
口腔保健衛生学専攻	IL2300597
口腔保健工学専攻	IL2300596
自由科目、複合領域コース	入門コース内のアンケートから回答してください。