

# DX推進のためのAI人材育成 研修サービス説明資料

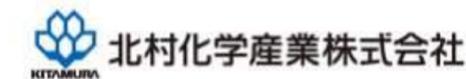
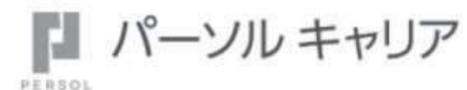
---



**AVILEN**

株式会社AVILEN

AI導入によって業界を牽引する大手企業様や急成長のベンチャー企業様など、業種業界を問わずご利用いただいております。(他多数)





## 代表取締役

データサイエンティスト  
崔 一鳴

東京工業大学大学院知能情報コース修了。日本語文書データ・画像データ・医薬データを対象にしたデータ分析・AIモデルの開発・コンサルティング経験が豊富。



## 取締役

ITスペシャリスト  
大川 遥平

筑波大学大学院システム情報工学研究科知能機能システム専攻博士前期過程修了。日本統計学会スポーツデータ分析コンペティション優秀賞受賞や各種学会にて発表。AI・Web・インフラと幅広くエンジニアチームをリード。



## CTO

データサイエンティスト  
吉田 拓真

東京大学工学部を卒業。図面読取AI・書類OCRシステムなどメイン開発プロジェクトをリードしてきた。世界最弱のオセロAIを開発し、多くのメディアにて紹介されている。



## 取締役

データサイエンティスト  
高橋 光太郎

東京大学大学院では機械学習の研究を行っていた。金融・製造・製薬などの業界を中心に、人工知能のビジネス適応を推進し、様々なプロジェクトでコンサルティング・開発に従事。日本ディープラーニング協会産業促進委員を務める。



## 取締役

ブランド責任者  
高橋 香輝

千葉大学大学院デザインコース修了。セールス、プロダクトデザイン・UI/UXデザイン・映像制作・WEBサービスの開発ディレクション等の領域を手がける。AVILENが関わる全領域のブランディングおよびデザイン制作の責任者を務める。



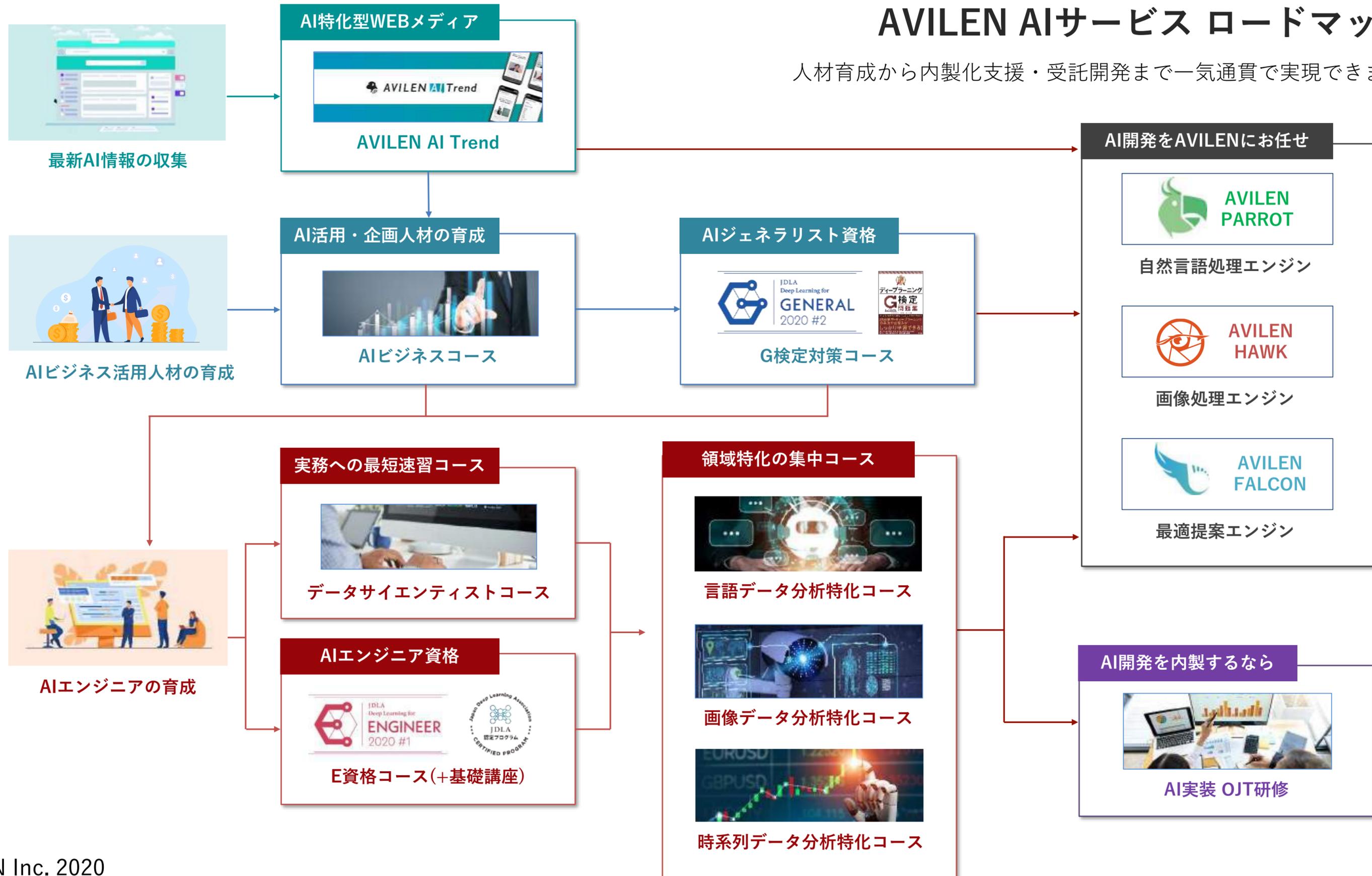
## 執行役員

データサイエンティスト  
吉川 武文

東京大学生物情報科学にてビッグデータの解析や生物学における理論のシミュレーション、モデリングなどにも精通。東京大学理科二類最高点合格、日本生物学オリンピック金賞・本選一位、合成生物学の世界大会iGEM金賞などの受賞歴を持つ。

# AVILEN AIサービス ロードマップ

人材育成から内製化支援・受託開発まで一気通貫で実現できます。



AI知識ゼロから始める

## AI導入・DX推進できる人材を育成

プログラミング不要

PoC成功率100%のAVILENが、AI導入ノウハウを教えます。

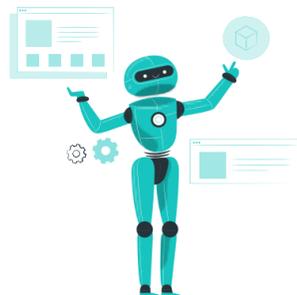
＼全てのビジネスパーソン向け／

＼DX/IT推進チーム・情報システム部門向け／

Level  
1

### AIビジネス**企画**コース

4.5 時間



AI・機械学習とは何か？  
基本のAIリテラシーを獲得



AIのビジネス活用を  
企画して要件定義できる

Level  
2

### AIビジネス**推進**コース

企画コース + 9.5 時間



様々な事例から、AI導入の  
手順や注意点がわかる



開発ベンダーと協力して  
AIプロジェクトを推進できる

## ①Eラーニング

動画視聴・WEBテストで基礎知識を習得



- ▶ 1チャプター8分程度の**動画講義を視聴**
- ▶ **小テストを解いて**理解度を確認
- ▶ メールの自動分類プロジェクトなど、事例を交えて解説

## ②課題提出（オンライン）

AIプロジェクトを整理するワーク



- ▶ 機械学習プロジェクトの要件定義を**実践**
- ▶ Eラーニングでの学習内容を整理
- ▶ 課題提出は**学習システムにアップロード**

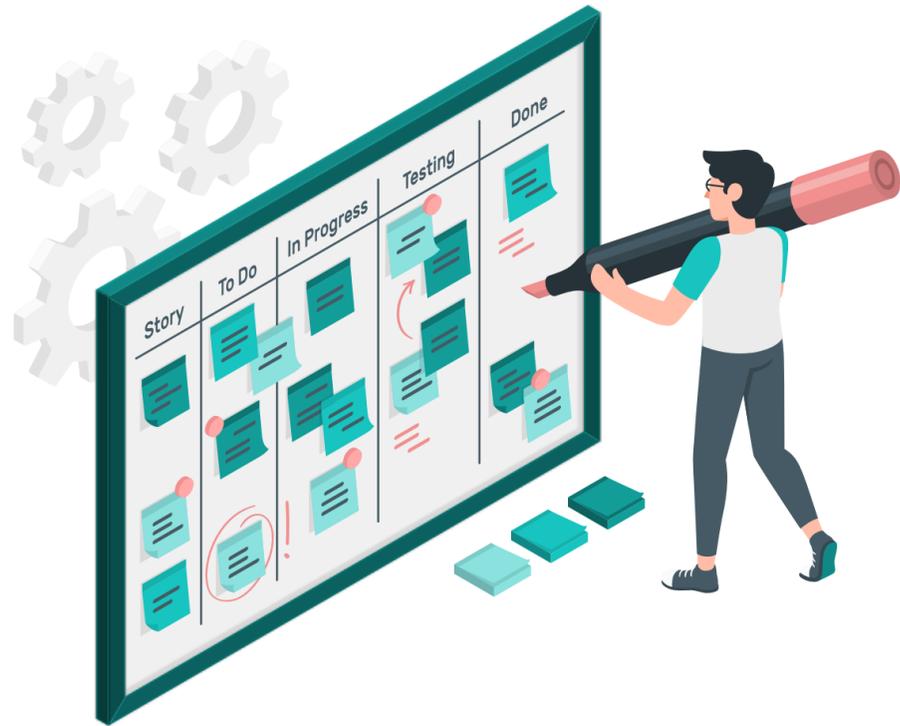
## ③ディスカッション会

全員参加型で講師より課題フィードバック



- ▶ Web会議形式で実施
- ▶ 講師から提出課題の内容を**フィードバック**
- ▶ 学んだ技術の応用について**ディスカッション**
- ▶ 積極的**アウトプット**により知識の定着を狙う

## AI・機械学習の基本や、 業務への活用企画を学ぶ



### AIとは・AIとの働き方

AIに多大な期待も過剰な期待もしてはいけません

#### 人間の脳と同等の機能を持つ機械はまだない

人間と同じように振る舞える汎用型AI

- 計画・予測など構造的に思考する
- 視覚
- 聴覚
- 嗅覚
- 触覚

流れに合わせた文章を読む

町屋や運動など思い通りに動く

© AVILEN Inc. 2020

### AIビジネスの全体像

AIビジネスの目的を達成する

#### 全体の流れ

© AVILEN Inc. 2020

### モデルの学習の流れ

メールの自動分類を例に

#### ① 訓練データで、モデルを学習させる

訓練データ → モデル1 → モデル2 → スパムメールの特徴を覚えた!

#### ② 検証データで、最良のモデルを選別する

検証データ → 学習済みモデル1 (92%) → 学習済みモデル2 (88%) → モデル1でテストしよう!

© AVILEN Inc. 2020

### AI導入の要件定義(初級)

実務ワーク

#### STEP① 目標の設定

良い目標とは?

- 様々な専門家同士が理解でき、コミュニケーションが行えるような目的や目標

例えば

- ○○に異常が起る48時間前までに予兆検知に成功する。
- ○○の目視チェックに掛かる人間の工数を90%以上削減する。

機械学習の特徴(精度100%はありえない)

- 技術上、必ず誤差を含むことを想定しておく。
- 目標となる精度は事前に想定して書きましょう。

© AVILEN Inc. 2020

## ベンダーとの交渉

### PoCで検証する範囲を決定

ベンダー: それぞれの流通量は具体的にどの程度でしょうか?

Ｃさん: メジャーな2種とマイナー8種は差がありますが、多いもので月800程度、少ないもので30程度です。

ベンダー: PoCではメジャー2種でAIの精度がどれくらい出るか見てみるのがいいかもしれません。

Ｃさんはこの発言を聞いて、「なぜ全部やらないのか」と疑問を持った。

© AVILEN Inc. 2020

## データの前処理方法

### マーケティング分野を通じて学ぶAIアプリケーション編 プロジェクトにおけるアノテーションイメージ

AIを訓練するために、データに答えをつけること

ID	名前	内容	アカウント	アノテーション
2019/12/27 16:38	宇野の低価格だったよ...	00000	必要	
2019/12/28 12:11	コスト削減のための良い！	00000	必要	
2019/12/29 19:53	サポートするよCCが当たる！	00000	必要	
2019/12/29 20:55	AMLEの議論は良いらしい	00000	必要	
2019/12/29 20:55	サポートするとCCが当たる！	00000	必要	
2020/1/1 16:19		00000		
2020/1/2 16:28		00000		
2020/1/3 18:38		00000		
2020/1/4 18:38		00000		

- ✓ 10000件のデータを必要か不要に分類する
- ✓ 少なくとも50-100時間以上かかる単純作業

© AVILEN Inc. 2020

# AIプロジェクトの具体的な進め方を学ぶ



## 著作権と著作権に注意

### X製薬も「共同著作物」に関する再契約を締結

X製薬とAIベンダーは、学習済みモデルを「共同著作物」とし、モデルの使用は「通訳AIの開発プロジェクトに限る」という取り決めを交わした。

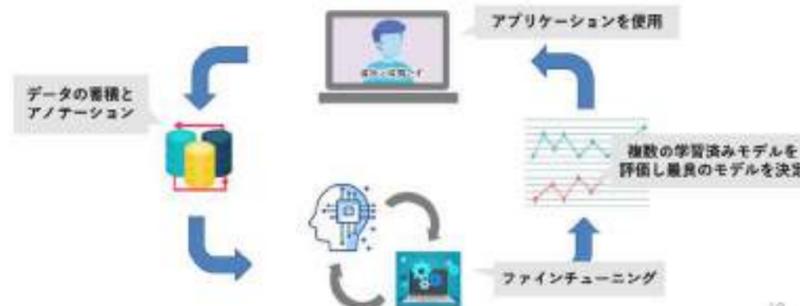


© AVILEN Inc. 2020

## アプリへの組み込み

### 使えば使うほど賢くなるAIを作る

データが自動的に溜まっていく仕組みをアプリケーションに組み込む



10

## G検定に合格し、AI知識を事業活用できるビジネス人材へ

AIジェネラリストを認定する”G検定”の合格に向け、AIに関する知識を網羅的に学びます。

### 最新の試験傾向に完全対応



2020年7月のG検定試験では  
**出題傾向が大幅に変化**



AVILENでは

**「Efficient Net」 「深層強化学習」**

等の最新分野まで詳細に解説！

＼G検定合格を目指すなら／

### G検定対策コース

30 時間



JDLA  
Deep Learning for  
**GENERAL**  
2020 #2

日本ディープラーニング協会 主催  
ジェネラリストG検定

- AI・機械学習・深層学習の理論
- ビジネス活用事例、法律
- 歴史から最新技術トレンドまで

＼5名以上の団体申し込みで／

### G検定問題集プレゼント

360 ページ



- 2020年9月発売の最新版問題集
- 丁寧でわかりやすい解説(170ページ)
- 重要ワードをまとめた用語解説つき

## ①Eラーニング

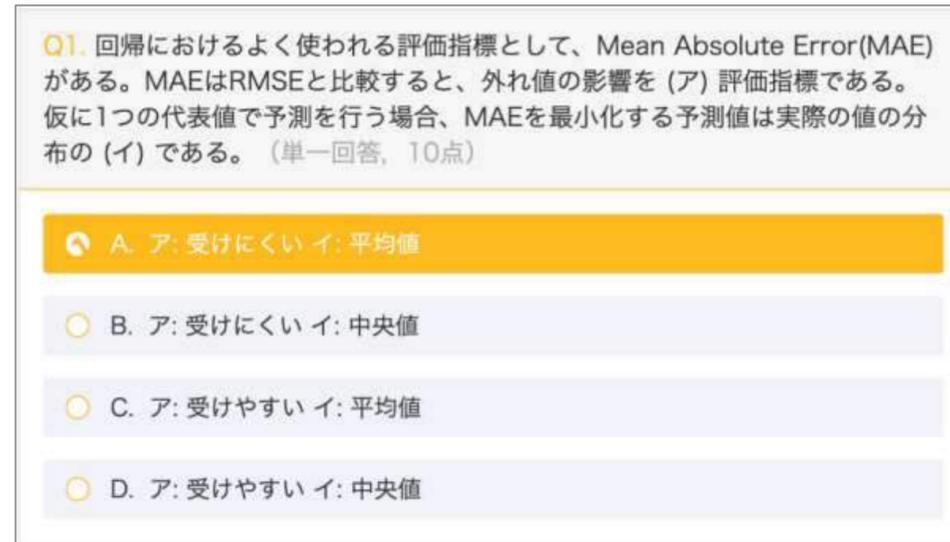
動画視聴しAI基礎知識を幅広く習得



- ▶ 専用テキストを元に**講師がわかりやすく解説**
- ▶ 章ごとに講義動画の**要点まとめシート**が用意してあり、試験前の復習などに活用できる

## ②WEBテスト (380問以上)

G検定と同じ4択問題で試験対策



- ▶ 380問の試験対策問題が**何度でも受け放題**
- ▶ 本試験と同じ4択形式のWEBテスト
- ▶ **即時に採点**され詳しい解説を閲覧できる
- ▶ 間違えた問題だけストックして復習も可能

## ③WEBテスト (模試)

G検定本番形式のオンライン模試



- ▶ G検定本番と同様の出題数、時間、4択問題
- ▶ 学習システム上から**いつでも受験可能**
- ▶ 問題は毎回ランダムで**何度でも受け放題**
- ▶ 本番での時間配分ミスなどを防ぐ

初心者から始める

## 最短期間でデータサイエンスの基礎を習得

実践的な開発演習

実務でのAI開発・運用に必要なスキルだけを厳選し、基礎を身につけます。

＼プログラミング未経験でもわかる／

＼業務に必要な最低限の実装スキル／

Level  
1

### 入門コース

18 時間



Python環境構築からサポート  
基本のデータ分析・機械学習



データベース操作(SQL)や  
AI活用事例も学べる

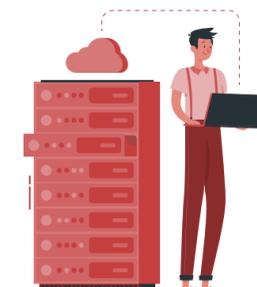
Level  
2

### 速習コース

入門コース + 34 時間

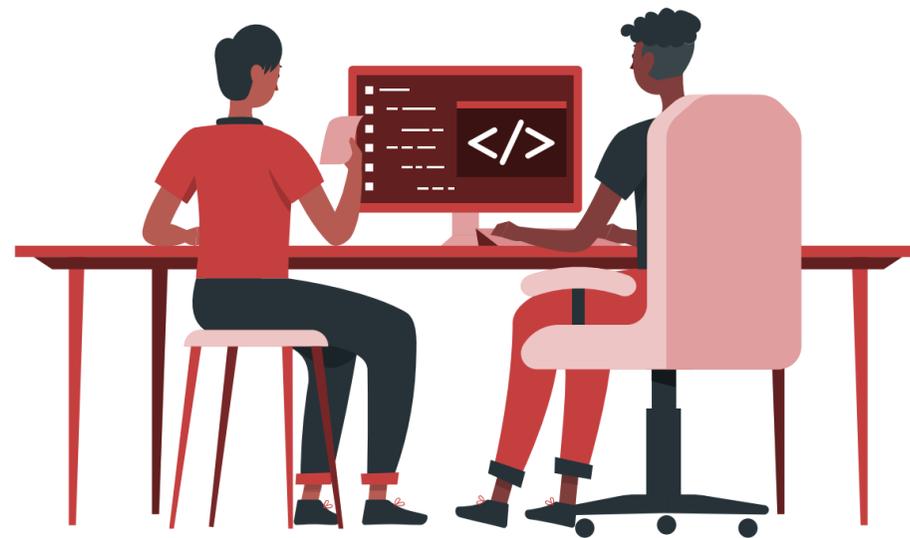


様々なPython・機械学習・  
ディープラーニング実装



モデル実装だけでなく  
インフラ知識まで習得

## Pythonの基本文法や、 機械学習の典型的なモデル の実装を学ぶ



### if/for文などの基本文法

実務上の  
型とは：どんな性質のデータをどう扱うか

例：数値・文字列どちらとして扱うかコンピューターに理解させたい



文字列  
50 → str型

数値10  
int型

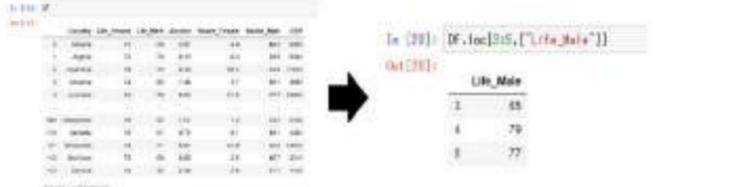
文字列"50"  
str型

違う型同士の計算はエラー

### Pandasでデータ整形

データフレーム入門  
範囲の抽出

・ `DF.loc[3:5,["Life_Male"]]`  
→ DFの3行目~5行目の"Life\_Male"の値を取り出せる！



読み込んだデータ全体

抽出したデータ

### scikit-learnで学習・検証

検証方法  
交差検証

データを分割し、複数回テストすることをk分割交差検証という。(クロスバリデーション、CVなどと呼ばれる)



全データ

ホールドアウト検証

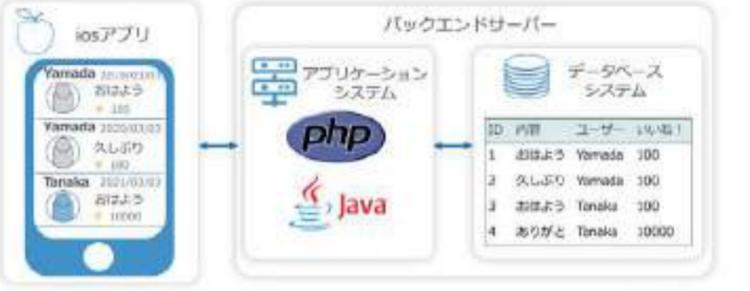
k分割交差検証

それぞれの結果の平均や、ぶれを確認

### データベース操作(SQL)

アプリケーションとデータベース  
クライアントとデータベース

クライアントは、一般的にはアプリケーションシステムを経由してデータベースサーバーにアクセス

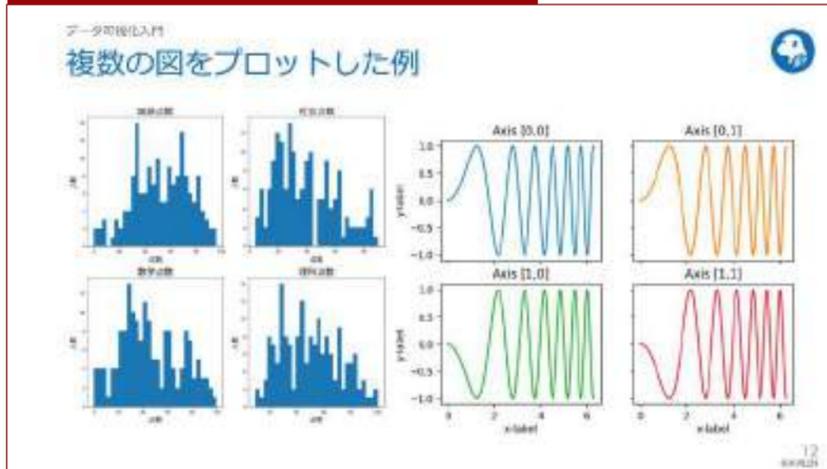


iosアプリ

バックエンドサーバー

データベースシステム

## NumPyやデータ可視化



## 機械学習の様々な手法を実装

### 【point】 高すぎる相関係数には要注意

一般的に説明変数を増やすほど決定係数が高くなりやすいため、分析者はついたくさんの説明変数を使ってしまう。しかし、相関が高すぎる値は機械学習モデルに対して悪影響を及ぼす可能性があります。(多重共線性) このような場合は、片方の説明変数を削除したほうが良いでしょう。

```
plt.figure(figsize=(20,20))
corr = imp_data.corr()
cmap = sns.color_palette("coolwarm", 200)
sns.heatmap(corr,xticklabels=corr.columns,
            yticklabels=corr.columns,
            square=True,
            annot=True,
            cmap=cmap)
plt.show()
```

# ディープラーニングや データ可視化、様々な 機械学習の実装を学ぶ



## PyTorchでディープラーニング実装

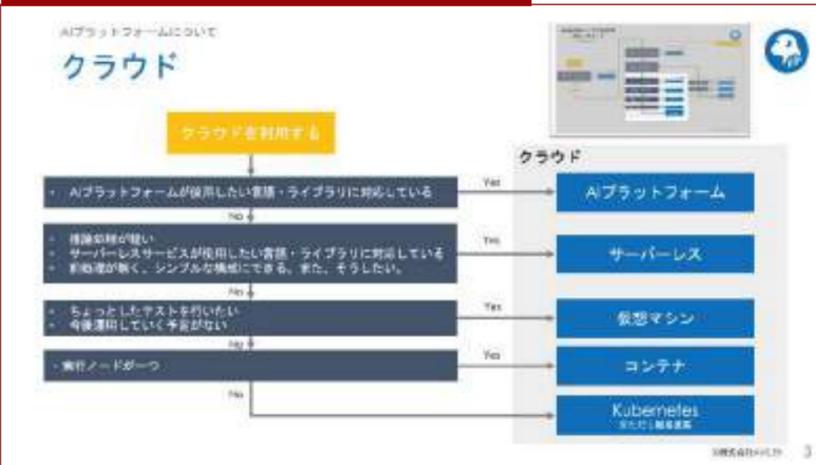
```
In [4]: class mlp_net(nn.Module):
def __init__(self):
    super().__init__()

    # 全結合層を2つ
    self.fc1 = nn.Linear(3, 5)
    self.fc2 = nn.Linear(5, 2)

    # 損失関数と最適化関数
    self.criterion = nn.MSELoss()
    self.optimizer = optim.Adam(self.parameters())

def forward(self, x):
    x = self.fc1(x)
    print('[fc1を通過]\n', x) # 中間層の出力を表示
    # 活性化関数
    x = F.relu(x)
    print('[reluを通過]\n', x) # relu()の適用結果を表示
    x = self.fc2(x)
    return x
```

## インフラの種類・選び方



基礎の後の応用編

## 専門性を深めたAIエンジニアを育成

実践的な開発演習

領域ごとに最新のディープラーニング手法を習得し、実務に活用できます。

＼時系列データ処理をマスター／

### 時系列データ特化コース

30 時間



- 時系列予測やレコメンドができる
- Prophet・SARIMA・LSTM・LightGBMなどの最新手法

＼画像処理をマスター／

### 画像データ特化コース

30 時間



- セグメンテーションや物体検出、異常検知、画像分類ができる
- CNN・ResNetなどの最新手法

＼自然言語処理をマスター／

### 言語データ特化コース

38 時間



- 文書分類や機械翻訳ができる
- BERT・word2vec・LSTM・Transformerなどの最新手法

## 時系列データ処理の基礎から応用まで網羅し 最新モデルの実装を学ぶ

こんなスキルが身に付く

- ✓ ベイズの定理など**統計学**の基礎
- ✓ **Prophet**時系列解析
- ✓ **SARIMA**・**状態空間**・**LightGBM**・**LSTM**  
による単変量時系列予測
- ✓ **SARIMAX**による多変量時系列予測
- ✓ 協調フィルタリングによる**レコメンド**

コーディング課題

### 店舗の売り上げ(需要)予測

データ：421,570行16列のテーブルデータ (.csv)



大手スーパーマーケット全店舗の毎週の売り上げと関連データから、各店舗の売り上げを予測するモデルを作成します。店舗ごとに予測の精度も出力できるようにします。

## 画像処理の基礎から応用まで 最新モデルの実装を学ぶ

こんなスキルが身に付く

- ✓ 画像分類に必要な**データ前処理**
- ✓ **ResNet**による画像分類
- ✓ **CNN**による画像認識
- ✓ **セグメンテーション** モデルの学習~評価
- ✓ **物体検出** モデルの学習~評価
- ✓ **異常検知** モデルの学習~評価

コーディング課題(中級)

### 花の画像から品種を予測

データ：17種類の低解像度カラー画像1,360枚

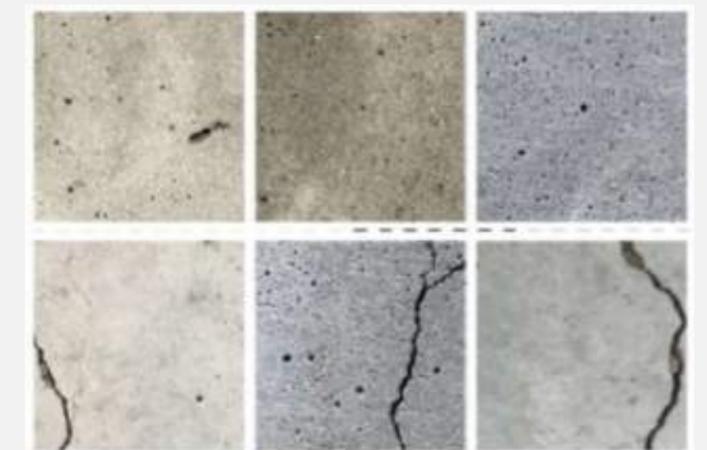


花屋で花の品種ごとに分類して値段をつけるシステムを構築するため、花の画像からその品種を予測できるモデルを開発していきます。

コーディング課題(上級)

### コンクリートの異常検知

データ：227×227pxのRGB画像40,000枚



建物のコンクリートのひび割れを自動で検知するシステムを作りたい。そこで、様々な建物のコンクリート画像から、異常を検知するAutoEncoderモデルを作成します。

## 言語処理の基礎から応用まで 最新モデルの実装を学ぶ

こんなスキルが身に付く

- ✓ 形態素解析・正規表現・文書のベクトル化
- ✓ word2vec理論と実装
- ✓ LSTMによる文書分類
- ✓ Transformer・Seq2Seqによる機械翻訳
- ✓ BiLSTM-CRFによる固有表現抽出
- ✓ BERTによる自然言語処理

コーディング課題(中級)

### ニュースサイト記事を自動分類

データ：8つのニュースサイトの7,376記事 (.text)



某大手ニュースサイトでは、記事の掲載先を人手で判別しています。コスト削減のため、未知のニュース記事に対して適切な掲載先を判別できるモデルを作成します。

コーディング課題(上級)

### 法律文書の自動分類

データ：10カテゴリーの法律文書235冊 (.text)



大量の法律文書データ (.text) を分野ごとに分類するタスクを自動化したい。そこで、未知の文書に対して適切な分類先を判別するディープラーニングモデルを作成します。

## 受講料金

**136,000**円 (税別)

## コンテンツ



講義 **動画** / 資料



**コーディング**演習 / 解説



プロダクト開発演習



**試験対策**問題



公式 **例題**解説動画

## サポート



チャット **質問**対応



課題コード **添削**



**学習相談** (ビデオ通話)

## 2020#1 E資格 合格率



- ・ 他社平均より約**20~30%**高い合格率
- ・ **100社**以上のクライアント&パートナー

## 2019#2 E資格 合格率



会社名	株式会社AVILEN (アヴィレン)
代表取締役	崔 一鳴
所在地	東京都中央区銀座2-14-4 銀座スクエア4F (本社)
スタッフ数 (2020年4月時点)	正社員：20名 アルバイト・業務委託：121名
事業内容	ディープラーニング技術を搭載したAIエンジンの開発・提供 AIの活用に関するコンサルティング支援 AI人材育成支援 AIメディア事業
所属組織	人工知能学会 日本ディープラーニング協会 データサイエンティスト協会
会社HP	<a href="https://avilen.co.jp/">https://avilen.co.jp/</a>

## 認定・認証



認証制度  
ISO(JIS Q)27001



JDLA認定プログラム  
認証番号：00008



JDLA公式メディア  
パートナー認定