

2020 年度青山アカデミア 講義概要 (シラバス)

2021 年 1 月 12 日

講義概要 目次

◆青山アカデメイア独自講座

* 統計・機械学習を利用したデータ分析入門（基礎編）・・・・・・・・ 1、2

* 統計・機械学習を利用したデータ分析入門（応用編）・・・・・・・・ 3

◆講座名	統計・機械学習を利用したデータ分析入門（基礎編）
◆開催日（回数）	全22回（e-learning12回、演習7回含む）
◆曜日	不定
◆時間	（講義）18時~19時（演習）18時~20時
◆会場（キャンパス）	オンライン形式
◆定員	5名

◆講座内容

データ分析に最低限必要な統計と確率の基礎、ならびに基本的な統計的手法・機械学習手法を用いたデータ分析について、e-learningによる座学とPythonというプログラミング言語を利用した演習を通して学ぶ。演習に利用するノートパソコンは講座期間中貸与する。

◆目標、重点を置く学習内容

難しい数学のテクニックを学ぶのではなく、実際に使える基本的なデータ分析の手法とその概念的な解釈を習得することを目標とする。

◆受講対象者

企業でデータ分析を新たに始める社会人、もしくはデータ分析の基本を再学習したい社会人を主な対象とする。ただし、Pythonを用いた演習を伴うため、コンピュータプログラムに対する一定の理解を有することが必要。

◆受講に際しての注意事項など

- ・最少催行人数を設けております。最少催行人数に達しなかった場合には、講座を中止させていただく場合がございます。
- ・講座の録音・録画・写真撮影は、ご遠慮ください。

◆講座スケジュール（各回の講義予定）

回	日程	内容
1	2月22日 (18-19時)	オリエンテーション 本講座の目的、学習内容の概要、進め方について説明する。
2	e-learning	データ分析の基本 データ分析の流れや要求される手法(技術)を簡単な分析事例を通して学ぶ。
3	e-learning	統計を利用したデータの読み方（1） 与えられたデータを読み解く道具立てとして代表的な統計量（平均、分散、標準偏差、中央値、最頻値など）について学ぶ。
4	e-learning	統計を利用したデータの読み方（2） 与えられたデータを読み解く道具立てとして度数分布グラフ、散布図などのデータの可視化方法に加え、2つのデータ系列間の関係を表す相関係数について学ぶ。
5	e-learning	データを読み解くための確率の基礎 データを読み解くために必要となる確率に関する基本事項を学ぶ。
6	3月1日 (18-19時)	理解度確認（1） これまでの学習項目についての理解度の確認と質疑対応を対面形式（オンライン）で行う。
7	e-learning	データからの統計量の推定 観察済みのデータから観察されていないデータを含む全データに対する統計量を推定する方法について学ぶ。
8	e-learning	統計的検定 データから導いた統計量の妥当性を統計的に判定する技術について学ぶ。
9	e-learning	単回帰分析 1つの数値データを別の単一の数値データを用いて説明する単回帰分析について学ぶ。
10	e-learning	重回帰分析 1つの数値データを別の複数の数値データを用いて説明する重回帰分析について学ぶ。
11	e-learning	機械学習の概要 機械学習とは何なのか、その処理の流れ、統計や確率に関する手法(技術)がどう使われるのかについて学ぶ。
12	e-learning	k最近傍法 最も単純な機械学習手法として、k最近傍法について学ぶ。

13	e-learning	決定木学習 カテゴリが付与されたデータから学習する決定木について学ぶ。
14	e-learning	クラスタリング 事前にカテゴリ（クラス）が付与されていないデータサンプルを類似するグループに振り分けるクラスタリングという技術について学ぶ。
15	3月12日 (18-19時)	理解度確認（2） これまでの学習項目についての理解度の確認と質疑対応を対面形式（オンライン）で行う。
16	3月15日 (18-20時)	（演習）Python の基本的な使い方 機械学習を学ぶ上で標準的なプログラミング言語となっているPythonの基本的な使い方、データの読み込み方法、便利に使える基本的なライブラリについてPythonを用いた演習形式で実践的に学ぶ。
17	3月16日 (18-20時)	（演習）Python によるデータの観察 Pythonを利用してデータから基本統計量を求める方法、およびデータを可視化する方法をPythonを用いた演習形式で実践的に学ぶ。
18	3月17日 (18-20時)	（演習）回帰分析 e-learningを通して学んだ単回帰分析、重回帰分析をPythonを用いた演習形式で実践的に学ぶ。
19	3月18日 (18-20時)	（演習）ロジスティック回帰 数値を予測する回帰分析の技術を利用してデータサンプルのカテゴリ（クラス）を予測するロジスティック回帰についてPythonを用いた演習形式で実践的に学ぶ。
20	3月19日 (18-20時)	（演習）決定木学習 e-learningを通して学んだ決定木学習をPythonを用いた演習形式で実践的に学ぶ。
21	3月22日 (18-20時)	（演習）アンサンブル学習 複数の決定木を組み合わせる機械学習手法についてPythonを用いた演習形式で実践的に学ぶ。
22	3月23日 (18-20時)	（演習）距離に基づく学習 e-learningを通して学んだk最近傍法とクラスタリングについてPythonを用いた演習形式で実践的に学ぶ。

◆講師紹介

大原 剛三：理工学部情報テクノロジー学科教授。博士（工学）。大阪大学産業科学研究所助手・助教を経て、2009年より青山学院大学理工学部情報テクノロジー学科准教授に着任。これまでに、人工知能学会理事、人工知能学会知識ベースシステム研究会主査・幹事を歴任。人工知能、特にデータマイニング、機械学習、社会ネットワーク分析の研究に従事。

小野田 崇：理工学部経営システム工学学科教授。博士（工学）。（一財）電力中央研究所を経て、2016年より青山学院大学理工学部経営システム工学学科教授に着任。これまでに、人工知能学会理事、品質管理学会製造業のためのビッグデータの解析あり方研究会主査を歴任。機械学習、特にニューラルネットワーク、サポートベクターマシン、機械学習に基づく設備診断の研究に従事。

◆講座名	統計・機械学習を利用したデータ分析入門（応用編）
◆開催日（回数）	全4回
◆曜日	不定
◆時間	18時~20時
◆会場（キャンパス）	オンライン形式
◆定員	5名

◆講座内容

統計・機械学習を利用したデータ分析入門（基礎編）で学んだ内容を踏まえ、より実践的にその技術を活用するために必要となるデータ前処理、日本語テキスト処理、および、深層学習の基礎をPythonというプログラミング言語を利用した演習を通して学ぶ。演習に利用するノートパソコンは講座期間中貸与する。

◆目標、重点を置く学習内容

個々の技術の詳細よりも、目的とする分析結果を得るためには、どのようなデータ処理が必要かを考える感覚を身につけることを目標とする。

◆受講対象者

統計・機械学習を利用したデータ分析入門（基礎編）を受講した社会人。

◆受講に際しての注意事項など

- ・最少催行人数を設けております。最少催行人数に達しなかった場合には、講座を中止させていただく場合がございます。
- ・講座の録音・録画・写真撮影は、ご遠慮ください。

◆講座スケジュール（各回の講義予定）

回	日程	内容
1	3月24日 (18-20時)	データ前処理 (1) 一部のデータ値が記録されていない欠損のあるデータの処理、および、膨大なデータ項目から必要な項目を選択する技術についてPythonを用いた演習形式で学ぶ。
2	3月25日 (18-20時)	データ前処理 (2) 膨大なデータ項目をもつデータをより少ないデータ項目のデータに変換する技術についてPythonを用いた演習形式で学ぶ。
3	3月26日 (18-20時)	テキスト分類 日本語テキストを複数のカテゴリに分類することを目的に、そのために必要となる文の分かち書き処理、単語頻度集計などの技術をPythonを用いた演習形式で学ぶ。
4	3月29日 (18-20時)	深層学習の基礎 深層学習の基礎を学び、既存の学習モデルを利用した画像分類をPythonを用いた演習形式で学ぶ。

◆講師紹介

大原 剛三：理工学部情報テクノロジー学科教授。博士（工学）。大阪大学産業科学研究所助手・助教を経て、2009年より青山学院大学理工学部情報テクノロジー学科准教授に着任。これまでに、人工知能学会理事、人工知能学会知識ベースシステム研究会主査・幹事を歴任。人工知能、特にデータマイニング、機械学習、社会ネットワーク分析の研究に従事。

小野田 崇：理工学部経営システム工学教授。博士（工学）。（一財）電力中央研究所を経て、2016年より青山学院大学理工学部経営システム工学科学科教授に着任。これまでに、人工知能学会理事、品質管理学会製造業のためのビッグデータの解析あり方研究会主査を歴任。機械学習、特にニューラルネットワーク、サポートベクターマシン、機械学習に基づく設備診断の研究に従事。