

YouTubeやブログでも学べる:

データサイエンティストを目指すなら知っておきたい学習方法

<https://techtarget.itmedia.co.jp/it/news/1902/22/news06.html>

データサイエンスを学ぶ場所と機会は大学の外側にも広がっている。大規模公開オンライン講座(MOOC)やオンラインで公開しているドキュメントなどがあり、データサイエンティストとしてのスキルセット獲得に役立つ。

2019年02月22日 05時00分 更新

[George Lawton, TechTarget]

関連キーワード

[MOOC](#) | [データ](#) | [データサイエンティスト](#) | [データ分析](#) | [スキル](#)

データサイエンティストのスキル向上と認定取得数増加を目的とする新しいプログラムが続々登場している。データサイエンティストはそこからさまざまなコースやトレーニングを選択できる。これまでも開講していたデータサイエンティスト向けの基本的なスキルセットを強化する大学のプログラムにも、新たなプログラムが追加されている。最新プログラムは、データサイエンスに必要なスキルをより強化する他、かつ他分野でも応用可能なスキルを習得できる内容となっている。



そのようなプログラムを受講してデータサイエンスの認定を取得すれば、データサイエンティストとしての市場価値を高められる。またデータサイエンスのスキルを応用できる機会が増え、豊富な経験値を蓄積できる。本稿では、データサイエンティスト向けの最新プログラムの中から、注目すべきものをいくつか紹介する。

併せて読みたいお薦め記事

データサイエンスに必要な基礎知識

- [データサイエンスチームに必要な“9つのプロフェッショナル”](#)
- [データサイエンティストを楽にする ビッグデータ分析ツールの10大「必須機能」](#)
- [いまさら聞けない「データサイエンティスト」と「ビジネスアナリスト」の違い](#)

機械学習を支援するツール類

- [機械学習プラットフォームでH2O.aiやKNIMEなど小規模ベンダーが躍進](#)
- [機械学習を使いやすくするAWSとGoogleのクラウドサービス、その利点は](#)

大学の講座で知識を深める

大学で開講しているデータサイエンスプログラムには、データサイエンスの基礎が詰め込まれているため、データサイエンティストとしてのスキルセットにおける自分の弱点を見極められる。基礎の他、データサイエンスを補完するツールや技法に関する情報も得られる。かつてこれらのプログラムは通学が必要で、修了には数年を要していた。しかし最新のデータサイエンスプログラムはオンラインで提供されるため、空き時間を使ってより効率的に学位を取得できる。

「『数週間でデータサイエンティストになる』というやり方は避けたい」。こう話すのは、データサイエンス企業Isimaの創設者、ダルシャン・ラワル氏だ。同氏は、学習範囲や習得期間が適正と思われるプログラムとして、米カリフォルニア大学バークレー校のデータサイエンス修士プログラムや、米カーネギーメロン大学のコンピューティングデータサイエンスプログラムなどを挙げる。

Velocity Groupのティム・ラファーティ氏によれば、米コロンビア大学のデータサイエンス認定プログラムもデータサイエンスに必要な分野を網羅しているという。このプログラムにはアルゴリズム、確率／統計、機械学習、視覚化に関する学習内容が含まれている。「このプログラムには開発における難易度の高い要素が盛り込まれている。これを受講することで、データサイエンティストは急拡大するデータサイエンスの業界において存在感を出せるようになる」(ラファーティ氏)

大規模公開オンライン講座

ここ数年、「Coursera」や「edX」といった、誰でも受講可能な大規模公開オンライン講座(MOOC)が続々登場している。MOOCは自身のペースでスキルアップを図れることが利点だ。無料のコースもあるが、データサイエンティストとしての認定を取得するには多少の費用が必要だ。

一例を挙げると、Courseraでアンドリュー・ウ氏が担当する機械学習コースでは、スタンフォード大学の認定を取得できるオプションがある。

「SuperDataScience」には、キリル・エレメンコ氏が担当する複数のオンラインコースがある。「エレメンコ氏は複雑な概念を4歳の私の娘にも理解できそうな方法で分かりやすく説明している。初心者にも、補修を必要としているデータサイエンティストにも役立つ内容だ」(ラファーティ氏)

短期集中型のデータサイエンスコース

オンラインコースではなく、講師やクラスメイトと実際に対面する短期集中講座を好むデータサイエンティストもいるだろう。例えば「Galvanize」は教室型のトレーニングプログラムで、初級クラスのデータサイエンティストを7週間で育成する。このプログラムは、実際のビジネスにおける問題にも向き合えるように、教室と企業との橋渡しをしている。

その他の短期集中コースとして、24週間で終了するカリフォルニア大学バークレー校の「Data Analytics Boot Camp」などがある。

スキルアップを目的とした非公式プログラム

ブログやソーシャルメディアからも、データサイエンスプログラムが幾つか誕生している。ケビン・マーカム氏が創設した「Data School」は、「YouTube」で受講できる。このプログラムでは数百時間にも及ぶデータサイエンスのコンテンツを提供している。

「アカデミックな学びを終えた後に学習を続けるためには、自分の属する業界で学びの機会を与えてくれる人物を探す必要がある」。こう語るのは、デジタルデザイン事務所SPRのパット・ライアン氏だ。同氏はマーカム氏のWebサイト訪問し、「Machine Learning with Text in Python」というプログラムに登録した。「マーカム氏は2018年7月に『Data School Insiders』を立ち上げている。これに毎月寄付することで、マーカム氏やフォーラム参加者への質問ができる」(ライアン氏)

ブログの「Towards Data Science」は、データサイエンスにおけるモデリングの概念を、初歩的なものから高度なものまで一通り解説している。人工知能(AI)やデータサイエンスに関する、専門家や研究者による最新の論文やプレゼンテーションへのリンクも紹介している。

ツールによる学習

「DataRobot」や「H2O Driverless AI」といった新しいデータサイエンスツールには、データサイエンスに必要なさまざまな要素を習得できるプログラムが含まれている。

「最も関心を寄せている新ツールは、グラフィカルで使いやすいUI(ユーザーインタフェース)を備え、データセットに最適なアルゴリズムを提供してくれるため、延々と試行錯誤を繰り返す必要がない。データサイエンスを普及させる可能性を秘めている」。こう話すのは、ソフトウェア会社Talendのイザベル・ヌアージュ氏だ。

このようなプログラムを含むツールはデータサイエンスのハードルを下げ、ビジネスアナリストなどが洗練された高度な分析手法を用いる手助けをする。

予測分析サービス会社Descartes Labsのエドゥアルド・フランコ氏によると、データサイエンティストが問題解決のため

に使用している手段を知るには、情報共有のためのプラットフォーム「Kaggle」が役立つという。ここには、多様なデータセットや実験的に活用できるプログラムの他、新しいスキルを習得するためのデータサイエンスプログラムなどが用意されている。

データサイエンスの課題と向き合う学習

データサイエンスの現場における新しい課題に向き合うことも、データサイエンティストのスキルを構築する方法の一つだ。

「まずは実際のデータサイエンスの課題を扱うのがお勧めだ。『習うより慣れろ』を実践して学べることは多い」。DatabricksのCEO（最高経営責任者）兼共同創設者のアリ・ゴージ氏はこう述べる。

データサイエンス以外に目を向けることも重要だ。

「数学者や宇宙論研究者、気候モデラー、コンピュータビジョンの専門家から学ぶのが私の日課だ」（フランコ氏）

こうしたデータサイエンスの学習機会は、「Insight Data Science Fellows Program」というデータサイエンス習得のための新プログラムでも取り入れられている。このプログラムでは7週間の課程を修了すると、受講生はデータサイエンスにおける新しい課題に取り組む機会を与えられる。

「受講生自身が希望する業界の入り口に立つとともに、専門家の力添えの下で実際のプロジェクトの経験を積める」（フランコ氏）

データサイエンスに関する無料の書籍で学習

最先端のデータサイエンス研究者や学者の一部が、データサイエンスに関する著書の全文をオンラインで無料公開している。これも「自分のペースで進められるデータサイエンスプログラム」の一種に位置付けられる。データサイエンスの新しい手法を学んだり、基本スキルを補強したりできる。以下に無料書籍の例を示す。

- 『Neural Networks and Deep Learning』、マイケル・ニールセン著
- 『An Introduction to Statistical Learning』、ガレス・ジェームス、ダニエラ・ウィッテン、トレバー・ハスティー、ロバート・ティブシャーアーニ著
- 『Pattern Recognition and Machine Learning』、クリストファー・ビショップ著

関連記事

[データサイエンスチームに必要な“9つのプロフェッショナル”](#)

[データサイエンティストを楽にする ビッグデータ分析ツールの10大「必須機能」](#)

Copyright © ITmedia, Inc. All Rights Reserved.

