

社会人向けデータサイエンス講座
独り立ちDS養成講座 & DS基礎講座
2024年度説明会

2024年03月05日

電気通信大学 データ教育センター

特任助教 佐野遼太郎

自己紹介

■ 名前

佐野遼太郎(Kaggle Master)

■ 仕事

電気通信大学特任助教

■ その他

競馬予測AI : Nimbaの開発・運用

競艇予測AI : さざ波の開発

仮想通貨Bot : Eurekaの開発

■ 経歴(データ分析技術者, データサイエンティスト)

✓ 2015年 3月: 名古屋大学 多元数理科学研究科 修了(微分幾何学)

✓ 2015年 4月: 株式会社金融エンジニアリング・グループ(金融・その他)

✓ 2018年 2月: 株式会社ディー・エヌ・エー(野球・危険運転)

✓ 2021年 4月: 富士通株式会社(器械体操)

✓ 2022年10月: 国立大学法人 電気通信大学(データサイエンス教育)

ガイダンス目次

- **本講座の概要**
 - 講座の概要
 - 講義概要
 - 受講に必要な準備
- **本講座の特徴**
 - 早くからのデータサイエンス教育の知見
 - 実践力重視
- **講義の概要**
 - 講義ラインナップ
 - 講義紹介
- **修了要件等**
 - 講義の履修について
 - 修了要件

➤ 本講座の概要

講座の狙いなど

➤ 狙い

- これから企業でデータ分析業務を行う人で、「見習い」レベルにある方を「独り立ち」レベルに引き上げます
- ビジネス経験は豊富だが今後の方向性に行き詰まりを感じている人に新しい可能性を付与します

➤ スタートレベル(受け入れの条件)

- 高校時代に数学(指数対数関数、三角関数、微分積分、ベクトル)を受けたか、または同等の内容をどこかで学習したことのある方
- プログラミング経験はあった方が良いですが、なくても構いません
 - その分、土日を勉強に当てる頑張ってもらうことになります

➤ 開催概要

- オンラインの講義とSlackを使った質疑応答(フル・オンライン)
- 土曜日の13時から18時(欠席者は録画資料を視聴していただけますが、それだけで修了要件をクリアできるわけではありません)

講座の概要

➤ データサイエンス教育の中での位置付け

- 基礎からエキスパートまでに対応するプログラム(DEFPP:データアントレプレナーフェロープログラム)から**必要なものを切り出した**
 - 「独り立ちデータサイエンティスト養成講座」:6ヶ月のコース
 - 「データサイエンス基礎講座」:3ヶ月のコース

➤ 学習の延長

- ご希望があり&必要な学力が相当と認められた受講生は、1つ上のクラスに必要な追加費用を払って、進むことができます
 - DS基礎 → 独り立ち:20万円+消費税
 - 独り立ち → エキスパート:30万円+消費税(特例であり早めの相談が必要)

※エキスパートの内容はDEFPPと同じですが、**修了証のタイトルが異なります**

※1年のうちに、DS基礎からエキスパートに進むことも可能です

- 学習の延長をしたいとお考えの際は、すぐに担当者(佐野)までご連絡ください。必要な学習について指示・指導いたします。
 - **確率・統計関係の学習について注意が必要**となります

UECデータサイエンス・プログラム

基礎レベル

コンピュータサイエンス

Pythonプログラミング

ITリテラシ

Python100本ノック

EDAとモデル作成の基礎

応用レベル

統計検定2級自習

回帰モデル作成

判別モデル作成

選択講義

エキスパートレベル

確率統計

DE実践論

特別講義

DS特論

教師なし学習

深層学習の基礎

自然言語処理

レコメンド

因果推論

文科省

リテラシーレベル

応用基礎plusレベル

エキスパートレベル

DS協会

見習いレベル

独り立ちレベル

棟梁レベル

到達目標

現在、データを活用して業務を行う上で必須とされるスキルを身につけた上で、データサイエンスのプロジェクトでリーダーに言われたことをきちんと実行できる。

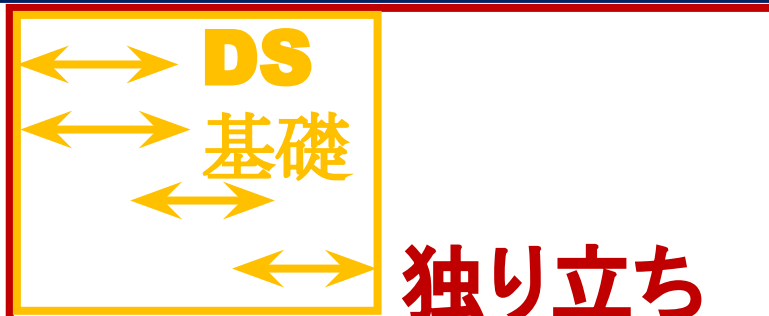
2つ3つ実践的PJを経験すれば、DSプロジェクトでリーダーの指示の下、必要なことを自力で実行することができる。

2,3年の実務経験の後、DSプロジェクトでリーダーとしてデータ活用の目的をクライアントと相談して策定し、相応しい分析を行った上で、結果を評価すること、また結果をシステム化するための要件を開発側に提示できる。

DSプログラム・開講スケジュール

4月 5月 6月 7月 8月 9月 10月 11月 12月 1月 2月

オンデマンド
・Pythonプログラミング
ITリテラシ
Python100本ノック
EDAとモデル作成の基礎



DS基礎:6月末まで
独り立ち:9月末まで

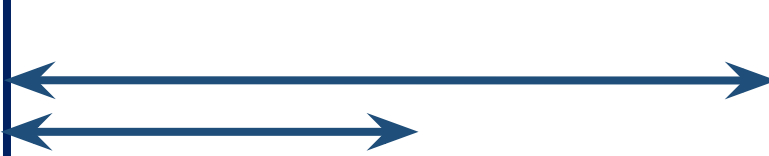
統計検定2級自習



回帰モデル作成
判別モデル作成



オンデマンド
・確率統計
・コンピュータサイエンス



DEFP

特別講義(必修)
DE実践論
Kaggle講義
DS特論



独り立ちでは選択科目

講義概要

➤ 講義形式

- オンラインによる遠隔講義のほか、eラーニングやオンデマンド講義を活用して、場所を選ばず、時間を有効活用できます
- Slackを活用することにより、講義時間以外でも講師との質疑や受講生同士での学び合いを行います

➤ 講義内容

- データサイエンスのリテラシーレベルの確認・復習から始め、最近の業務環境での作業方法を身につけていただきます
- データサイエンスプロジェクトに必要なプログラミング力、サイエンス力を学びます
- 実データをできるだけ用いることにより、実際のビジネスにデータサイエンスを展開する実践力を身につけます

➤ 学習条件

- 講義の日程は土曜日に設定していますが、業務時間外での**自習が必須**です
- 業務時間外での学習にも講師陣はSlackで対応いたします

受講に必要な準備(ネットワークやマシンについて)

➤ マシンの要件

会社のPCだとインストールやアクセスに制限がある場合があるので、**会社のPCは非推奨**

- WindowsでもMacでも構いませんが、平日、休日、業務時間外でも持ち帰って学習できるようなノートPCをご用意ください
 - MacBookAirでも良いですが、HDは512GB以上をお勧めします
 - メモリは16GB以上を推奨します
- Pythonの統合環境「Anaconda」をインストールして使っていただきます

➤ ネットワーク

- Zoomによるオンライン会議システムに入れるネットワークが必要です
- 質疑等のコミュニケーションをSlackで行いますので、Slackが使えるネットワークが必要です
- 計算環境としてGoogle Colaboratory, GCPにアクセスできる必要があります

➤ 会議環境

- 1社から複数名ご参加の場合でも、会議室に集まってまとめてのご参加ではなく、個人個人でご参加ください

➤ クレジットカード

- 実質無料ですが、GCPの利用にクレジットカード番号が必要です

➤ 本講座の特徴など

ITリテラシ教育によるDX対応

➤ データサイエンス開始以前の問題

- ビッグデータの時代、データや情報を扱う業務は、これまでとは仕事のやりかたや、お作法が全く違ってきました
 - データはクラウドにあるから「SQL書いて自分でとってきて」
 - 計算もクラウドでやるから、「GoogleのアカウントでColaboratoryを使って」
 - プログラムのバージョン管理は「GitHubでね」
- こうした「新しい仕事のやり方」を身につける必要があります

➤ これまでの実績で積み重ねた経験を注入した講義

- これまでの受講生がどこで詰まったのか、何がハードルとなっているのかを熟知した担当者が、受講生を指導した経験を教材化
 - cf. 受講生の声:「職場の若い人とのコミュニケーションがスムーズになりました。これまではSlackだのGitHubだのと言った若いメンバーの会話に入れなかったのですが、自分でも会話に入るだけでなく、実践できるので、若い人たちから「おおっ！」って驚かれています、前よりもコミュニケーションが取れて、とてもやりやすくなりました」

電通大におけるDS教育の蓄積

➤ 早くからのデータサイエンス講座の実施経験

- 2015年から寄附講座・大学院の特別講義として大学院生および社会人向けのデータサイエンス講座を開講
 - 滋賀大学のデータサイエンス学部の設立が 2016年
 - データサイエンティスト協会の「DS養成講座」開講が2016年5月

➤ 多くの実地経験に基づく教育ノウハウ

- 基礎学習をeラーニングで行うこと
- 実際の企業データを借りてきて演習を行うこと
- ビジネス課題を実際のデータから導き出すこと
- Kaggleの積極的活用

実践力=実際に「使える」力の重視

➤ 電通大の伝統

- 無線電信講習所以来、「先端の科学を社会に実装する」ことを重視
- 口がよく動く人よりも、**手がよく動く人材**を育てます
- 単に知見が広まりましたということだけでなく、会社に帰って、**価値を生み出すことができる**ようになります

➤ データアントレプレナーフェロー・プログラムの知見を活用する

- いち早くPBL(Project Based Learning: 課題解決型学習)を採用
 - DEFPでは、アスクル様、DeNA様、小田急電鉄様、ゴルフダイジェスト・オンライン様等の皆様にデータスポンサーになっていただいています

到達目標:DS基礎

➤ 見習いにエントリー可能:DS部署に配属されても大丈夫

- DSプロジェクトのメンバーとすることができます
 - PLから言われたことが出来るようになります
 - サーバからSQLでデータを取ってきたり、クラウド上で指示通りの集計ができた
り、ベースラインのモデルを指示されたアルゴリズムで実行できるようになります
- 習得スキル
 - 現在の先端の当たり前、情報やデータを使っての仕事の仕方が身につきます
 - コード(Python)を書いて、前処理、可視化を行うことができます

➤ 各種スキル・レベル

- Python3エンジニア認定基礎試験相当
 - 修了生のほとんどは、受験すれば合格できると思います

到達目標: 独り立ちDS養成

➤ 独り立ち: DS人材として「貢献できる」人材に

回帰と判別のモデル作成の
ポイントを学べます！

- 見習いレベルの終了: 独り立ちレベルへ
 - PLから言われたら、集計、可視化、モデル作りをこなすことができます
- 習得スキル
 - 現在の先端の当たり前、情報やデータを使っての仕事の仕方が身につきます
 - コード(Python)を書いて、前処理、可視化を行うことができます
 - 機械学習のモデルを作ることができるだけでなく、ビジネス的側面をも考慮した総合的なモデルの評価を行うことができるようになります

➤ 各種スキル・レベル

- 統計検定2級相当
 - 修了生のほとんどには試験を受けて合格してもらいます
- Python3エンジニア認定基礎試験相当
 - 修了生のほとんどは、受験すれば合格できると思います

➤ 講義科目の紹介

講義科目 (以下は状況によって配信方法などが変更になる可能性があります)

➤ オンデマンド:必修科目

- Pythonプログラミング: データサイエンスのプログラミングを学びます
- コンピューターサイエンス: IT技術の基礎となるコンピュータについての理論を学びます
- 統計検定2級合格支援: **独り立ちのみ**

➤ 遠隔講義:オンライン講義

- ITリテラシ
- 特別講義: Python100本ノック
- EDAとモデル作成の基礎
- 判別モデル作成: **独り立ちのみ**
- 回帰モデル作成: **独り立ちのみ**
- デザイン思考: **独り立ちのみ**
- データサイエンティストによるモック(MVP)作成: **独り立ちのみ**

必修科目

選択科目だが受講を推奨!

選択科目

➤ オンデマンド(学習資料と映像による補助教材): 選択科目(**独り立ちのみ**)

- 自然言語処理, 因果推論, 深層学習の基礎, レコメンドモデル作成, 教師なし学習の基礎と演習

講義概要:ITリテラシ

➤ コンピューターで仕事をする

- コンピュータの構造:Linuxの基本コマンド
- クラウド環境での作業:GCEの活用
- 文字:文字コード、正規表現

➤ ネットワーク環境で仕事をする

- TCP/IP入門
- セキュリティの基礎
- HTMLとWebサイト

➤ チームで仕事をする

- GitとGitHub
- コミュニケーションツールの活用

➤ 大規模データの取り扱い

- データベース概論
- SQLの基礎

講義概要：100本ノック&EDAとモデル作成の基礎

➤ Python100本ノック

- データサイエンティスト協会の「データサイエンス100本ノック・構造化データ加工編のPython版」をやります
- 各個人が100問できるまでやります

➤ EDAとモデル作成の基礎

- 集計と可視化の力をしっかりつけます
- 機械学習モデルの目的や限界を学びます
 - 現代の高度な機械学習モデルの実用において最も重要なポイントの1つである「検証 (Validation)」の意義について、学習します
- 教師あり学習を実際にプログラミングしながら学びます
 - 判別モデル: titanic
 - 回帰モデル: Jリーグの観客数動員予測

➤ 目を瞑っていても、集計と可視化はできるようになって下さい

実践的な機械学習モデル作成講義(独り立ちのみ)

➤ 学習内容

- 機械学習の基本となるモデル
 - 回帰モデル: 不動産価格予測(担当: G.A.Technologies)
 - 判別モデル: マーケティングデータ分析(担当: コニカミノルタ)

➤ 使用データ

- Kaggleのデータおよびコンソーシアム参加企業から提供いただいたデータ
 - 全て実務で実際に使われていたデータですから、机上の空論ではありません

➤ 講師陣

- 現在も第一線で活躍しているデータサイエンティストが、単に精度のよいモデルを作るのではなく、実際に「使える」モデルの作成を指南します
 - データ取り扱いの実践的ノウハウ
 - 業務知識をデータ分析やモデル作成にどう活かすのか

➤ 修了要件など

講義の履修について

➤ 講義とお休み

- 期間中毎週土曜日の13時～18時までの講義
- **講義のお休みは2024年5月4日(ゴールデンウィーク)のみ**
- 独り立ちの方々は、8月24日以降は「仕上げ」期間として、オンデマンド講義の受講と進捗度別の個別対応(主に統計検定2級)を行います
- 4月6日の午後からガイダンス(対面orオンライン)があります

➤ 講義への出席

- 各講義は録画しているので、参加できなかった場合でも視聴することはできます
- しかし、講座の後半にある講義ではグループ・ワークを行いますし、講義の中では理解を確かめるために講師からの質問等、インタラクティブな授業をやっていきますので、録画を見ているだけでは十分とは言えません
 - 講義の理解度、進捗度は講義出席率に比例するようです。録画を見ているだけでは修了レベルに達することは難しいとお考え下さい

➤ 自習の必要

- 予習の必要はありませんが、**復習は必要**です
- 講義を受けている時間の2倍から3倍の時間の復習が必要です

修了と修了要件

➤ 修了証の授与

- 下記の要件を満たした方には修了証を発行します
- この修了は、何らかの資格の取得を意味するものではありません

➤ 修了要件: 以下の全ての条件を満たしている必要があります

- 期間中に必修のオンデマンド講義を受講して、その中で指定されている課題を提出し、合格と認められること
- 独り立ちの方は、期間中に「統計検定2級」の試験に合格すること
- 独り立ちの方は、回帰モデル作成、判別モデル作成の講義に出席して、グループワーク等の中で必要とされるスキルを発揮して、モデル作成を行うこと

費用

➤ DS基礎講座

- 30万円＋消費税

➤ 独り立ちDS養成講座

- 50万円＋消費税

➤ 学生の扱い

- これらの講座には学生無料の枠はありません。例え、現役の学生であっても、費用をお支払い下さい

募集締切日 & 卒業生コミュニティ

➤ 2024年3月15日(金)締切

- 応募書類等は下記ページをご参照ください
 - ◆ <https://de.uec.ac.jp/zero/2024entry/>

皆様のご応募をお待ちしております！

➤ 卒業生のコミュニティもあり, 年2回の勉強会等の実施想定

- 基本的には卒業生が登壇(内容はDS関連であれば自由)
 - ◆ データ分析業務での困りごと共有
 - ◆ 勉強した内容の発表
 - ◆ Kaggleへの取り組み内容発表