

# DS基礎講座 2022秋コース 説明会

2022年9月7日

電気通信大学

齊藤 史朗

# ガイダンス Agenda

## ➤ 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

- ✓ スケジュール
- ✓ 修了要件

## ➤ **本講座の概要**

# 講座の概要

## ➤ 狙い

- ✓ これから企業でデータや情報を使った業務に携わる人に、現在におけるデータを使った業務のやり方の基礎を身につけていただきます。
- ✓ 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。

## ➤ 対象レベル

- ✓ 高校時代に数学II (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
- ✓ 年齢不問

## ➤ 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時まで。(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。録画視聴はないよりましとお考え下さい)。

## ➤ 期間と費用など

- ✓ 2021年11月から1月 (3ヶ月) :2023年4月～6月のコースもあり。
- ✓ 費用：一人当たり30万円+税 (お支払いいただいた費用は、途中離脱などの場合の、ご返金には応じかねますので、ご了承ください。)
- ✓ 修了証を発行します。

# 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、サイエンス力の基礎を学びます。

## ➤ 学習条件

- ✓ 講義の日程は土曜日に設定していますが、平日の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で対応いたしますので、ぜひ、土日・休日、業務時間後も時間を作って学習してください。

# 参加に必要なもの（ネットワークやマシンについて）

## ➤ マシンの要件

- ✓ 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。
  - もしこの講習用に新たにPCを購入なさるのであればMacのノートブックをご購入ください。（Unix環境が簡単に手に入るのです。）
  - MacBookAirでも良いですが、HDは512G以上をお勧めします。
- ✓ 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。
  - メモリ8G以上(16G以上推奨)、空きディスクが200G以上が必要です。

## ➤ 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。
- ✓ 共通した計算環境を得るためにGoogle Colaboratoryにもアクセスできるネットワークをご用意ください。

## ➤ 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。
  - 指名して画面共有していただいて、音声で質問を受けたり、指示をいたしますので、マイクやスピーカーは個人に属している必要があります。

## ➤ **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方が 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
- ✓ Pythonは基本的なことができるまでやってもらいます
- ✓ Kaggleの積極的活用：コンペの環境を使ってモデル作成の基本を身につけていただきます

# 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る。
- ✓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、「データ活用」ができるようになります。

## ➤データアントレプレナー・プログラムの知見を活用する

- ✓多くの社会人に教えてきて、何が足りないのか、何を身につけないと現代のビジネスシーンで活躍できないのかの知見をためてきた
  - 本心にDXをするために
  - データを活用する業務環境に慣れる

# 到達目標

## ➤ DS人材として「貢献できる」人材に

- ✓ データサイエンス業務をやったことのない人を「見習いレベル」にします
  - 言われたことが「理解」でき、DS業務の「働き方」に適応できます
- ✓ 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とは、どう言うことなのかを理解できるようになります

## ➤ 各種スキル・レベル

- ✓ Python3エンジニア認定基礎試験相当
  - 集計や可視化は「目を瞑ってでもできる」ことを目指します
- ✓ モデル作成
  - 判別モデルおよび回帰モデルについてコンペに投稿できるレベルになります

## ➤ 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。)

## ➤ Eラーニング：必修科目（オンデマンド）

- ✓ 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ーについての理論を学びます。

## ➤ 遠隔講義：必修科目（オンライン講義）

### ✓ ITリテラシ

- Linuxの使い方
- クラウド環境で仕事をする
- GitHubによるバージョン管理
- データベースの基礎

### ✓ python100本ノック

- DS協会の「python100本ノック」をやって、できるようになってもらいます

### ✓ EDAとモデル作成の基礎

- データの集計と可視化
- 判別モデルと回帰モデルの作成とコンペサイトへの投稿

# 講義概要（抜粋）：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

# 実践的な機械学習モデル作成講義

## ➤ 学習内容

- ✓ まずは集計と可視化：モデル作成の基礎
- ✓ 機械学習の基本となるモデル
  - 判別モデル
  - 回帰モデル

## ➤ 使用データ

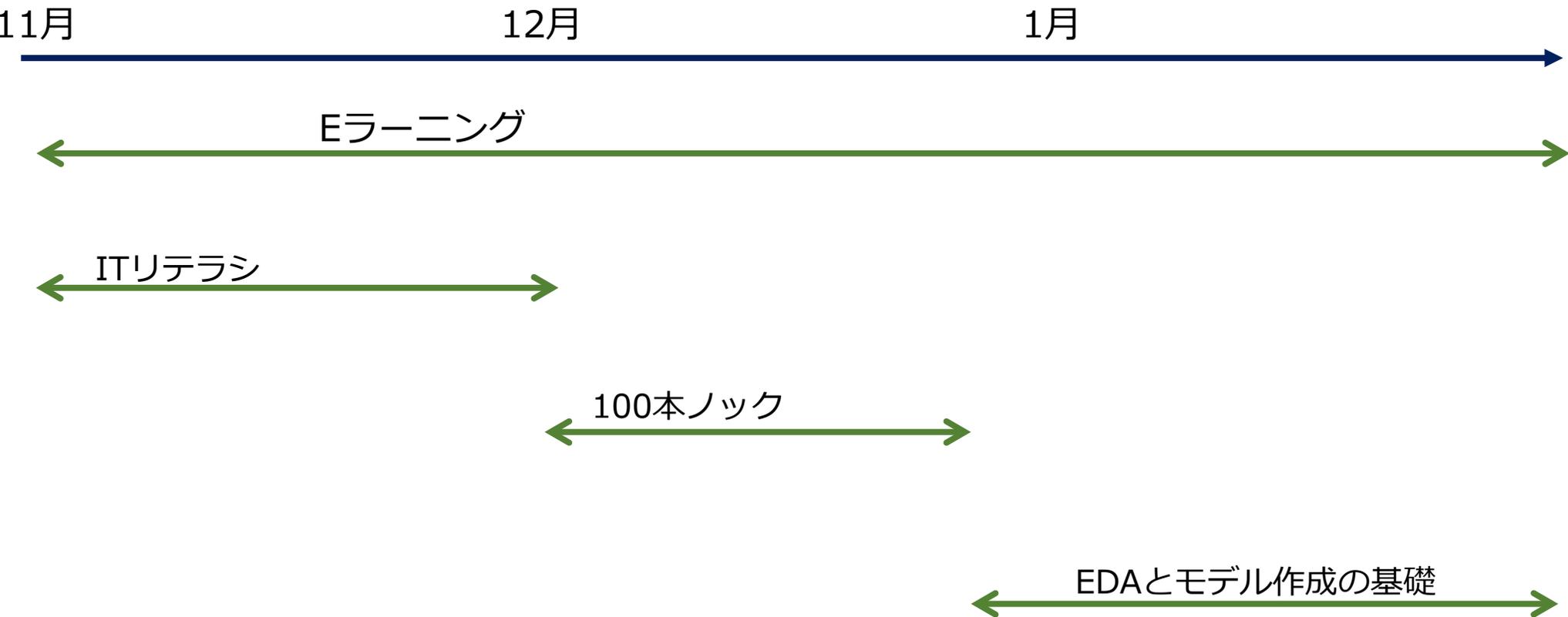
- ✓ KaggleやSignateのコンペサイトの練習問題データ
  - コンペの形でモデルの精度向上を目指すことによって、モデル作成に必要なスキルを身につけます

## ➤ 講師

- ✓ 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
  - 精度評価の意味

➤ **講義スケジュールや修了要件など**

# スケジュール



# 講義の履修について

## ➤ 講義とお休み

- ✓ 期間中（11月5日から1月28日）毎週土曜日の13時～18時までの講義。
- ✓ 講義のお休みは12月31日のみ。

## ➤ 講義への出席

- ✓ 各講義は録画してありますので、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終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。

- ✓ 1月末までにeラーニングを受講して、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。
- ✓ 「EDAとモデル作成の基礎」講義の中で行うコンペサイトへの投稿結果が講義で指示する水準を超えていること
  - 「ITリテラシ」については、終了要件を作りませんが、この講義で身につけるべきスキルが身に付いていないと、その後の講義に参加すること自体が難しくなります。
  - 「Python100本ノック」はできるまでやってもらいますが、ここで学んだことがしっかり身に付いていないと、「EDAとモデル作成の基礎」講義で所定の期限までに水準を超えるモデルを作成することはできません。