

NTT DATA

NTT DATA Mathematical Systems Inc.

新バージョン発表会【データサイエンス講座】

NTTデータ数理システム

Trusted Global Innovator

NTT DATA Group

NTT DATA

新バージョン発表会【データサイエンス講座】

1. データサイエンスの現状
2. データサイエンス講座プログラムの概要
3. データサイエンス講座で用いる分析ソフトウェア
4. 講座の費用、今後の日程

データサイエンスの現状

企業のデータサイエンスへの取り組み状況

世の中の企業の取り組みの現状は？

- データサイエンティストの囲い込み
- 新規にDX推進部門の立ち上げ
- AI民主化の加速による、各種分析サービスの活用

実際のお悩みは？

- データサイエンティストの人材難
- 社内に有識者がいないので、進め方が分からない。
- 費用対効果がはっきりしないので、手を出しにくい。

自社でデータサイエンスに取り組む手段

| 手段 | メリット | デメリット |
|------------------------|--|---|
| ①データサイエンスを行っている企業に外注する | <ul style="list-style-type: none">予算さえ工面すれば素早く結果を出すことができる | <ul style="list-style-type: none">外注する際に要件を適切に取りまとめたり、必要なデータを蓄積し提供するスキルは必要。やってみないと効果が判断できないこともある |
| ②データサイエンティストを新規に雇う | <ul style="list-style-type: none">良い人が募集できれば手っ取り早い | <ul style="list-style-type: none">かなりの売り手市場なので募集しても集まらない。高給、使う側にもそれなりに知識が必要。 |
| ③データサイエンティストを自社で育成する | <ul style="list-style-type: none">データサイエンスに関する知見を自社内で蓄積することができる。 | <ul style="list-style-type: none">一定の成果を出すまでに、それなりの時間がかかる。 |

③データサイエンティストを自社で育成する---にアプローチする講座のご紹介

データサイエンス 講座プログラムの概要

データサイエンス力向上に向けたプログラムです

※2014年12月10日データサイエンティスト協会様プレスリリースより引用

データサイエンス力
(分析結果から結論を導く)

データエンジニアリング力
(統計、プログラム)

ビジネス力
(課題の整理)

本セミナーでは【データサイエンス力】に関する教育のお手伝いをします。

【データエンジニアリング力】 ⇒ 当社データ分析ツールの利用で補完

受講後の目標

1. 一般的なAIの知識と活用方法について理解する
2. 様々な手法に対して、基本的な分析作業ができる
3. 与えられた問題設定に対して、分析手法を選択し、設計や分析計画を行い、実行。業務に役立つ分析結果を得る
4. 実務上の課題に対して問題設定を行い、分析を依頼、分析者が作成した結果を実務に活用できる

場所を選ばないWebinar開催です！

講義内容（各論では操作実習を行います）

| | 講義名 | 内容 |
|-----------------|----------------|--|
| 概論 | AI概論 | AIとは？核となるデータサイエンスの関連技術やプロジェクトの進め方を紹介。マネージャーの方にもお薦め |
| 各論 1 | 課題設定とデータ分析の考え方 | 課題設定の心得、課題解決に必要なデータとは？ データの種類と分析手法の関係 |
| | 前処理、基礎分析と可視化 | 集計、欠損の扱い、グラフ表示、アソシエーション分析 |
| 各論 2 | 予測のための線形モデル | 線形回帰、ロジスティック回帰 |
| | データ構造の把握 | クラスタリング、多変量解析（主成分分析を中心に） |
| 各論 3 | 予測のための非線形モデル | 決定木、Neural Network、Random Forest、 勾配ブースティング、Deep Learning |
| | モデルの定量的評価と最適化 | 交差検証、評価指標(ROC、AUCなど) モデル選択とハイパーパラメータ最適化 |
| 各論 4 (オプション) | 時系列データの分析 | 時系列データの前処理、基礎分析、 線形モデルや機械学習による予測、分類 |
| 各論 5 (オプション) | テキストマイニング概論 | 自然言語処理、テキストデータの分類 その他テキストマイニング分析手法 |

過去の講座の担当講師



アルゴリズム開発からデータ分析など機械学習全般に関する受託分析に従事。近年は異常検知に関するパッケージ開発やコンサルティング、データサイエンス教育プログラムなどを中心に取り組む。2020年度のデータサイエンス教育プログラムの各論講師を担当。



自然言語処理や機械学習全般に関する受託分析やパッケージ開発、コンサルティングに従事。近年は異常検知に関するパッケージ開発やコンサルティング、データサイエンス教育プログラムなどを中心に取り組む。2019年度、2020年度のデータサイエンス教育プログラムの各論講師を担当。

Webinar講習の様相(QAの様子)

『AI概論』におけるQAの例

【質問】

AI検証やPoCのプロジェクトの進め方について、最初に業務設計を明確化し決められた分析作業を実施するのか、それとも、ある一定の期間内で臨機応変にご要望に応えながら作業するのか、どちらの場合が多いのでしょうか？

【回答】

基本的なゴール（目指す先や大まかな問題設定）は分析に入る前に合意を取り決めていて、その中で詳細な分析や前処理の方法は分析者が考えることが多いです。但し、途中で先方の要望を聞いて多少方向転換をすることもあります。

『予測のための非線形モデル』におけるQAの例

【質問】

接点の不純度についてどのようなパラメータで、どのような値を設定すべきなのか（演習では0.0007なのはなぜか？）教えてください。

【回答】

分岐基準により変わるが0.01以下を選択すると無難化と思います。実際にはパラメータ探索を行って決めることが多いです。

今までのご参加者様からいただいたご感想

講座内容について

- データサイエンスの重要なエッセンスについて一通り学ぶことができ、大変勉強になりました。
- データ解析手法および評価方法を学ぶことができ、業務に役立てることができそうです。
- 実業務でまだ未経験の手法も多かったので、大変勉強になりました。
- 本講座を受講することにより、データサイエンス力を向上することができ、大変嬉しく思います。
- 難しく煩雑な分析がとても楽にできると感じました。

講座資料について

- 資料が充実していて、復習しやすいと思いました。また補足資料も大変分かり易かったです。
- 分析設計(各手法のメリットデメリットやパラメータ選択の際の注意事項など)に関する説明が分かり易く参考になりました。
- モデル選定やパラメータ最適化は、業務でも悩む部分であるため、このように資料にまとめ、解説までしてもらえて本当に助かりました。

オンライン学習について

- 分からないことは全て回答してくれるので、安心して臨めました。
- オンラインでの講座はトラブルが予想されますが、トラブルなくスムーズに進んで良かったです。
- 社会情勢も厳しいなか、こんなにしっかりと研修をWEBから受講でき大変良かったです。

Webページに掲載中です！

<https://www.msi.co.jp/solutions/DS/testimonials.html>

データサイエンス講座の特徴・メリット

実践的なデータサイエンス教材

当社が長年培った統計・データ分析のノウハウを凝集させたオリジナルのテキストを使用します。データサイエンスに必要な基本項目から、より実践的な分析方法、最先端のデータサイエンス領域までを短時間で学習することができます。

ユーザーフレンドリーな解説

難解な数式の表記は必要最低限とし、分かりやすい解説に努めています。

(数式に関するご説明は補足資料としてご用意しています。)

実例を交えた演習

データ分析に関する一連の流れについて、実例に近いデータをもとに演習を行い理解を深めます。

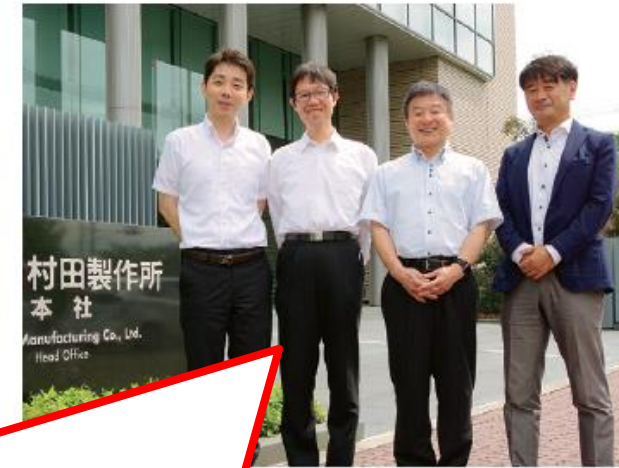
分析経験豊富な講師陣

当社の最前線で豊富にデータ分析を経験している講師陣が、みなさまを強力にサポートいたします。

株式会社村田製作所 共通基盤技術センター 様

AIを全社で身につける

—データサイエンティスト育成プログラムを実施—



一般教養を身に付けるようにAI技術を習得
させたい、という意図で、
プログラム

前から
NTT

デー
ラムを

NTTデータ数理システムのツールを使用すれば業務が効率化し、
仮説立案など本来業務に集中できるようになります。
機械学習モデルの構築や検証に**VMS**、統計モデル構築のために**VRP**、
さらにディープラーニングモデル構築に**Deep Leaner**を使っています。

(注から51人目)

<https://www.msi.co.jp/solutions/DS/index.html>

データサイエンス講座で 用いる 分析ソフトウェア



データ活用の確かなパートナー

なぜ、RやPythonでの講習ではないか？

習得にかかる時間の削減、分析のスピードアップ

Rだと・・・無料だけど・・・

```
R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R は、自由なソフトウェアであり、「完全に無保証」です。
一定の条件に従えば、自由にこれを再配布することができます。
配布条件の詳細に関しては、'license()' あるいは 'licence()' と入力してくださ

R は多くの貢献者による共同プロジェクトです。
詳しくは 'contributors()' と入力してください。
また、R や R のパッケージを出版物で引用する際の形式については
'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが出ます。
'help.start()' で HTML ブラウザによるヘルプがみられます。
'q()' と入力すれば R を終了します。

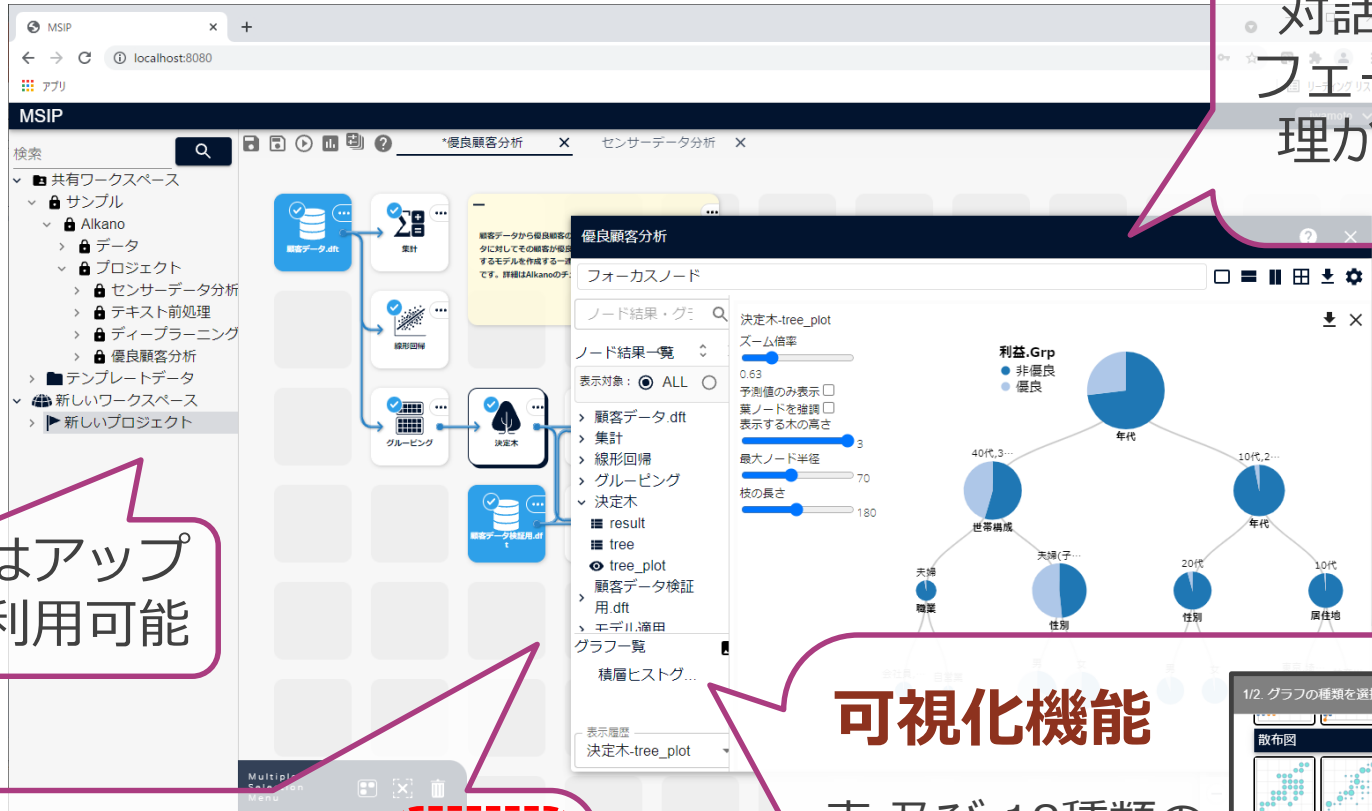
[以前にセーブされたワークスペースを復帰します]

> a <- 1:10
> |
```

?これ?
これで本当に分析ができるの？

プログラム不要、データサイエンスに専念して学習できます

対話的なインターフェースでデータ処理から、分析まで



外部データはアップロードで即利用可能

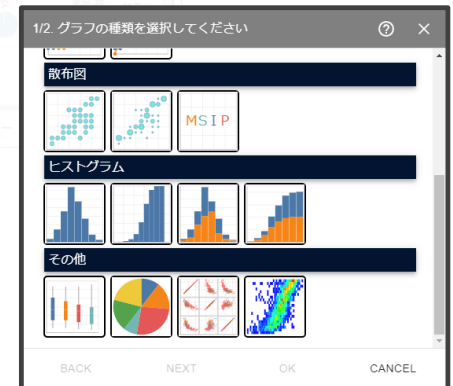
処理の流れをフローで構築



機能の単位：アイコン

可視化機能

表及び18種類のグラフによりデータや分析結果を確認



講座の費用 今後の開催日程

開催日程

2021年度第4回の公開講座のスケジュールは下記のとおりです。

| | 講義名 | 日時 | |
|-----------------|--------------------------------|-----------|-------------|
| 概論 | AI概論 | 1月12日 (水) | 13:30~16:00 |
| 各論 1 | 課題設定とデータ分析の考え方 前処理、基礎分析と可視化 | 1月19日 (水) | 10:00~16:00 |
| 各論 2 | 予測のための線形モデル データ構造の把握 | 1月26日 (水) | 10:00~16:00 |
| 各論 3 | 予測のための非線形モデル モデルの定量的評価と最適化 | 2月 2日 (水) | 10:00~16:00 |
| 各論 4 (オプション) | 時系列データの分析 | 2月 9日 (水) | 10:00~12:30 |
| 各論 5 (オプション) | テキストマイニング概論 | 2月 9日 (水) | 13:30~16:00 |

費用

| | 講義名 | AI関連技術 |
|--------------|----------------|---------------------------------|
| 概論 | AI概論 | — |
| 各論 1 | 課題設定とデータ分析の考え方 | データの種類、モデル設計 |
| | 前処理、基礎分析と可視化 | 前処理、基礎分析、可視化 |
| 各論 2 | 予測のための線形モデル | 線形回帰、ロジスティック回帰 |
| | データ構造の把握 | 多変量解析、クラスタリング、次元圧縮 |
| 各論 3 | 予測のための非線形モデル | 決定木、Random Forest、Deep Learning |
| | モデルの定量的評価と最適化 | モデル評価、ハイパーパラメータ最適化 |
| 各論 4 (オプション) | 時系列データの分析 | 基礎分析、欠損補間、ARIMA、Deep Learning |
| 各論 5 (オプション) | テキストマイニング概論 | 自然言語処理、クラスタリング、分類モデル |

| 価格 | 基本 (概論、各論 1～3) | 基本+各論 4 | 基本+各論 5 | 全て (概論、各論 1～5) |
|------------|-------------------|----------|----------|-------------------|
| 一般受講者様向け価格 | ¥282,900 | ¥321,900 | ¥330,700 | ¥369,700 |
| ユーザー様向け価格 | ¥226,250 | ¥257,500 | ¥264,500 | ¥295,750 |

価格には2か月分のデータサイエンス講座用ライセンス(Alkano)が含まれます。各論5には2か月分のText Mining Studioライセンスが含まれます。業務でのご利用はご遠慮ください。

※金額は税抜です。

データ分析サポート（オプション）

データサイエンス講座で得られた知識を業務に活かす際、疑問に思われたこと、実際にデータを前にされて困った点などに対応します（データサイエンス講座の受講は必須ではありません。本サポートのみのご発注にも対応しますが、予め、基礎的な知識は身に付けて望まれたほうが効率が良いです）

| | |
|------|--|
| | 5時間(NTTデータ数理システムのソフトウェアを用いた、リモート会議システムによるサポート) |
| 内容 | <ul style="list-style-type: none">・ 課題、データ、分析最終目標などのヒアリング・ サポート終了時の目標設定・ データ傾向把握、理解のための分析<ul style="list-style-type: none">・ データ読み込み / 集計 / 数値、カテゴリーの加工 / 可視化・ 目標に向けた、分析の選択<ul style="list-style-type: none">・ データ内容、目標に応じた手法の選択方法・ 分析実行、結果の把握<ul style="list-style-type: none">・ 分析の実行、評価・ 結果の説明、理解のための質疑応答・ 今後の方針について、ディスカッション<ul style="list-style-type: none">・ 今後どのような分析が有効か・ どのようなデータが必要か |
| 見積金額 | ・ ¥300,000-(5時間、Web会議にて2回に分けての実施が可能です) |
| | <ul style="list-style-type: none">- 貴社ご担当者様との事前打ち合わせ、あるいは事前調査、分析などは含みません（データ受け渡し、電話やメールでの短い打ち合わせは想定しています）- ご参加者様は4名様までです- 当社ツールを使ったサポートに限ります- 納品物は特にございませませんが、作成した分析プロジェクトや、メモをお渡ししています |

sales@ml.msi.co.jp まで
メールにてご連絡ください！

件名：「データサイエンス講座」見積依頼

貴社名：

お名前：

ご連絡先：

株式会社 NTTデータ数理システム

TEL : 03 - 3358 - 6681

【URL】 <https://www.msi.co.jp>

【E-mail】 sales@ml.msi.co.jp