



日本初 製造業特化型データサイエンス人材育成

T.B.TECH
テービーテック株式会社

製造業のデータサイエンスの 現状・課題



製造業のデータサイエンスの現状・課題

人材不足

社内・工場内の
AIへの認知の温度差

データの
不足・不備

本講座からのアプローチ

人材不足

自社でドメイン知識のあるデータサイエンティストをそろえたいが、教育には時間がかかる・・・通常業務に影響が出ることは控えたい

AI開発は外部に委託しているが、提案されたことが最適解なのか判断できる人材がいらない

忙しい社会人の最大のネックは **時間**

昼の部と夜の部の講義を用意し、業務スタイルに合った時間帯で受講可能な環境を整備

講義内容を製造業に特化することで、業務に必要な知識を効率的に吸収できるカリキュラムを用意

同じ製造業に携わる受講生が集まるので知見の共有が容易

業務スタイルにあった受講時間と必要な知識の選別・濃縮で時間的な縛りを緩和し効率よく業務に反映することで受講のハードルを下げました

求められる人材ひとつ **橋渡し役**

体系的に製造業に必要なAIの知識を得ることでブラックボックス化を防止

AI開発の中で何が行われているかを理解し、必要・不必要の判断が可能になる
また、自社の課題や要望を正しく伝えられるようになります

講座中に他社との交流の中で様々な事例を聞くことができ、自社に置き換えた場合など考察を深めることができます

本講座からのアプローチ

社内・工場内のAIへの認知の温度差

AIに対する認識の違いから過度な期待・落胆が生まれるため、PoCから先に進むことが難しい
(現場が求めるのは100%の精度だが、現実には不可能)

現場の職人の理解を得られないことがある

データ収集のためやAIの解析結果を反映させるために製造ラインなどに手を加える・変化を与えることに現場の担当者などがためらう



データサイエンティストの神髄は **コミュニケーション**

製造業のAI活用は現場との連携が不可欠です
AI解析の精度を上げるだけでなく
AIの解析結果から得られる利益・リスクの認識を正しく伝え、共有することが求められます

本講座では演習のコードレビュー・ディスカッション・PBLなどのグループワークを豊富に取り入れることで

他者の意見を聞く、自身の意見を伝える

それらを**統括**して最良の結論を出す力が身につきます
また、ひとつの意見を押し通すだけでなく、
条件によって柔軟に手段を講じ、提案する力をやしません



本講座からのアプローチ

データ不足・不備

現状がアナログデータの蓄積が多く、データの規格などが不揃い

異常検知にAIを活用したいが異常データがそもそも少ない

データはあるが、課題を検証することに適したデータではないもしくは足りない



データが無いからできない、で終わらない
提案力

課題を解決するためにどのようなデータがあることが理想で現在手元にあるデータからはどの程度の成果が見込めるか判断する力が現場では必要になります

データへのアプローチは多種多様
対処方法の引き出しを増やすことが重要です

演習やPBLを通して様々な手法や事例を学び、データクレンジング・不均衡データへのアプローチなど必要なデータをどう収集すればいいかの道筋を構築し、データが無いからできない、で停滞せず次の案を提案する力を身に付けます

「製造業特化型 データサイエンス集中コース」

「製造業特化型データサイエンス集中コース」 概要

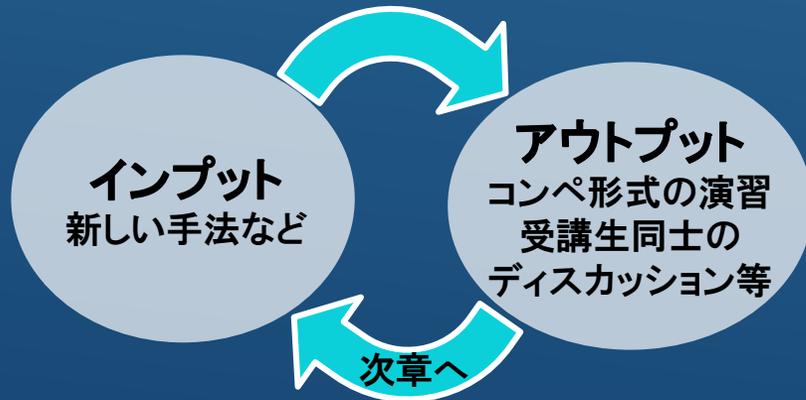
目的

「現場」視点で考えることのできる データサイエンティストの育成

一般的なデータサイエンス養成コースでは、幅広い内容を学ぶことができる一方、各業界固有の問題に特化して学ぶことはできません

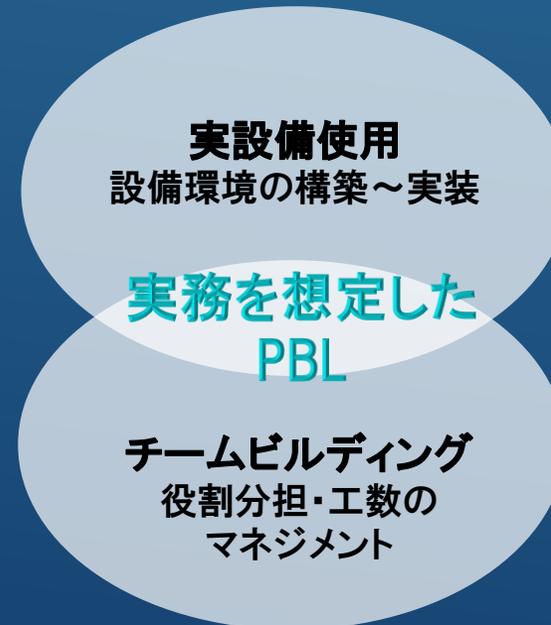
本コースは多変量解析、画像処理、時系列解析、最適化といった製造業の現場で本当に必要とされる領域に絞って集中的に学ぶことができます

前半3ヶ月



各章ごとにインプットとアウトプットを繰り返す教科書通りの答えだけではなく、自ら調べて独自に工夫を凝らしたアウトプットを受講生同士で共有し、知見を広げていきます

後半3ヶ月



前半3ヶ月で培った知識と経験を活かしてチーム戦を行います
実設備を使用し、期間・費用などの条件のもとデータの収集～モデルの実装・結果のアウトプットまでの実務により近い発想のPBLとなっております

受講して得られるスキル

現場の問題を定式化して コンピュータに解かせる力

在庫の最適化や人員計画のスケジューリングなど、現場にある問題をいかにして、コンピュータに解かせるための定式化を行えるかを考えられるようになります
詳細な数学も大切ですが、本コースでは「現場で使える」を優先したカリキュラムを用意しています

幅広い領域に関する 知見を習得

「AI」と一言で表しても、多変量解析や画像処理、時系列解析など様々な領域を含んでいます
本コースでは、製造業で必要となる領域をピックアップしているため、本当に必要な知識を効率よく幅広く習得することができます

クラウドを含めたIoTまで 考慮したシステム構築力

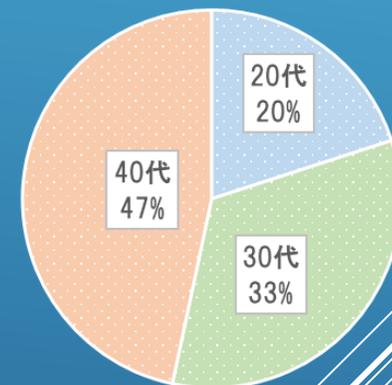
工場内の製造機器のデータをネットワーク経由で取得・分析し、故障予知や品質管理を実現すれば、企業にとって大きなメリットが生まれます
その一方で、AIだけでなくクラウドを含めたIoTまでを考慮した知見を得るためには幅広い知識と経験が必要となります
本コースでは、このシステム構築にも重きを置いています

幅広い層に対応したカリキュラム

※円グラフは豊田校1期の受講生アンケートより

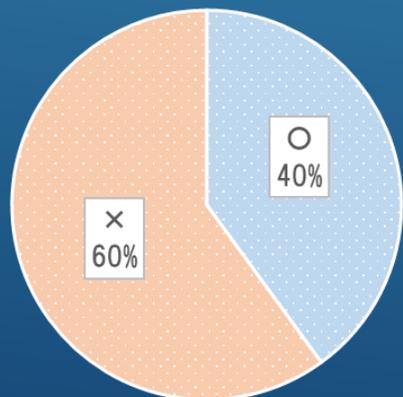
年齢

20代～40代と幅広い世代の方にお集まりいただきました
それぞれ社内での立場の違い、経験の違いから発見の多い充実した議論が可能となりました



普段の業務で既に機械学習に取り組んでいるか？

初心者の方は今後の業務等に向けての導入として、
既に業務で機械学習に取り組んでいる方は、
独学であいまいになってしまっている部分を体系的に学ぶことを目的としていました



製造業に特化したインプット

事前学習動画

9時間程度

イントロダクション

人工知能・機械学種・ディープラーニングとは
機械学習に必要な数学
機械学習の3大トピック
内挿と外挿

微分

微分は「何」に使えるのか？
導関数を求めよう
微分の公式
偏微分

単回帰分析（数学）

「モデル」を決める
「評価関数」を決める
評価関数を「最小化」する

Python入門

プログラミングの環境構築
変数
基本構文
複数の変数を扱う
制御構文
関数

単回帰分析（実装）

Numpyの数値計算
Pandasによるデータベース操作
Matplotlibでグラフの描画
実データに対する単回帰分析の実装

線形代数

スカラー・ベクトル・行列
行列の演算
サイズ感
転置・単位行列・逆行列
ベクトルで微分

重回帰分析（数学）

「モデル」を求める
「評価関数」を決める
評価関数を「最初化」する

重回帰分析（実装）

行列演算の基礎
パラメータの導出
実データで演習

統計

主な統計量
正規分布と3 σ 法
スケーリング

外れ値を考慮した実装

外れ値除去
モデル構築
スケーリングとパラメータの確認

ビジネス活用

現場で機械学習を導入できる人材とは

製造業に特化したインプット

Month1

Day1 イン트로ダクション

- 1.1 人工知能・機械学習・ディープラーニングとは
 - 1.2 学習と推論
 - 1.3 機械学習の導入事例
 - 1.4 機械学習プロジェクトに必要な人材
 - 1.5 AIプロジェクトのフロー
 - 1.6 演習
- 参考: AIシステムの開発フロー
- 1.7 機械学習案件に必要となる人材とは

Day2 データ解析の基礎

- 2.1. Google Colaboratory
- 2.2. Pythonの基礎
- 2.3. pandasの基礎
- 2.4. matplotlibの基礎
- 2.5. scikit-learnの基礎

Day3 代表的な回帰の手法と前処理 1

- 3.1. scikit-learn 入門
- 3.2. 各ステップの改善
- 3.3. 代表的な回帰の手法 1
- 3.4. 相関関係と多重共線性問題
- 3.5. 代表的な前処理 1

Day4 代表的な回帰の手法と前処理 2

- 4.1. 代表的な回帰の手法 2
- 4.2. 演習
- 4.2. 演習解説
- 4.5. 分類の問題設定に挑戦
- 4.6. 演習2

Day5 ハイパーパラメータのチューニングと分類の代表的手法

- 5.1. ハイパーパラメータのチューニング
- 5.2. ハイパーパラメータのチューニング手法
- 5.3. 代表的な分類の手法
- 5.4. (参考)教師なし学習

Day6 モデルの評価方法と演習

- 6.1. 分類のモデルの評価指標
- 6.2. scikit-learnでそれぞれの指標の確認
- 6.3. 不均衡データへのアプローチ
- 6.4. その他の分類の評価指標
- 6.4.1. 正規化された混同行列
- 6.4.2. ROC曲線とAUC
- 6.5. 演習

Day7 テーブルデータ演習

- 7.1. コンペ内容
- 7.2. データセット&評価指標
- 7.3. ベースラインとなるモデルの作成
- 7.4. EDA(Exploratory Data Analysis)

Day8 ニューラルネットワークの基礎

- 8.1. ニューラルネットワークとは
- 8.2. ニューラルネットワークの計算
- 8.3. ニューラルネットワークの訓練

製造業に特化したインプット

Month2

Day9 TensorFlow入門

- 9.1. TensorFlowとは？
- 9.2. Kerasの基礎
- 9.3. Kerasでクラス分類
学習済みモデルを保存
学習済みモデルを使用した推論
- 9.4. 一連の流れを確認
- 9.5. 入力変数と教師データの作成
演習課題

Day10. CNN (Convolutional Neural Network)

- 10.1. 問題設定
- 10.2. MNISTで練習
- 10.3. CNNモデルの定義
学習結果の確認
- 10.5. Function APIの使い方
- 10.6. 学習済みモデルの保存
- 10.7. 学習済みモデルによる推論

Day11. 演習：画像分類

Day12. 転移学習と ファインチューニング1

- 12.1. データセットの準備
- 12.2. 入力変数と教師データの作成
- 12.3. 訓練データと検証データの分割
- 12.4. シンプルなモデルで学習
- 12.5. 転移学習とファインチューニング
- 12.6. モデル構造の可視化&保存
- 12.7. 演習課題

Day12-2 転移学習と ファインチューニング2

- 12.1. データセットの準備
- 12.2. 入力変数と教師データの作成
- 12.3. 訓練データと検証データの分割
- 12.4. シンプルなモデルで学習
- 12.5. 転移学習とファインチューニング
12.5.1. ファインチューニング(Fine Tuning)

Day13. CNNの精度を向上させる テクニック

- 13.0. ベースモデルの構築
- 13.1. Data Augmentation
- 13.2. 最適化アルゴリズム
- 13.3. 過学習対策
- 13.4. アンサンブル学習
- 13.5. (参考)活性化関数

Day14. 物体検出

- 14.1. 物体検出の歴史
14.1.0. 早速使ってみよう！
- 14.2. 物体検出と画像認識の違い
- 14.3. 物体検出の流れ
- 14.4. モデルの性能を測るための評価指標
- 14.5-1. 代表的な物体検出手法
- 14.5-2. YOLOv3実装

Day15. 生成モデルによる異常検知

- 15.1. 導入
- 15.2. 異常検知モデルの学習方法
- 15.3. 異常検知で用いられるモデル
15.3-1. AutoEncoder
15.3-2. VAE(Variational Autoencoder)
15.3-3. ALOCC
(Adversarially Learned One-Class Classifier)

製造業に特化したインプット

Month3

Day16. セマンティック セグメンテーション

- 16.1. イントロダクション
- 16.2. データ作成方法
- 16.3. FCN(Fully Convolutional Networks)
- 16.4. U-Net

Day17. 画像認識演習

- 17.1. 画像分類
- 17.2. 異常検知

Day18. 時系列データ (RNN)

- 18.1 時系列解析の基礎
 - 18.2. 時系列解析で分類
 - 18.3. RNN(Reccurent Neural Network)
 - 18.4. 演習課題(60分)
- モデルへのアプローチ : CNN

Day19. 異常検知 (Anomaly Detection)

- 19.1. 異常の種類
- 19.2 身の回りの異常を考える
- 19.3. 異常検知の手法
- 19.3-1. ホテリングT2法
- 19.3-2. OneClassSVM

Day20. 異常検知 (Anomaly Detection)

- 異常検知の手順
- 20.1. 異常部位検知
- 20.2. 変化点検知
- 20.3. その他の時系列処理

Day21. 画像認識演習

Day22. 時系列演習

Day23 エッジデバイス

- 23.1. Raspberry Pi とは?
- 23.2. Raspberry Piの各種設定
- 23.2.1. SDカードの用意
- 23.2.2. SDカードをラズパイに装着し起動
- 23.2.3. IPアドレスの固定
- 23.2.4. システムアップデート
- 23.2.5. 必要モジュールのインストール
- 23.2.6. Python3関連

Day24. Web アプリケーション

- 24.0. Web の基礎
- 24.1. Flaskの基礎
- 24.2. Hello Worldを表示
- 24.3. FlaskでWebAPIを作成

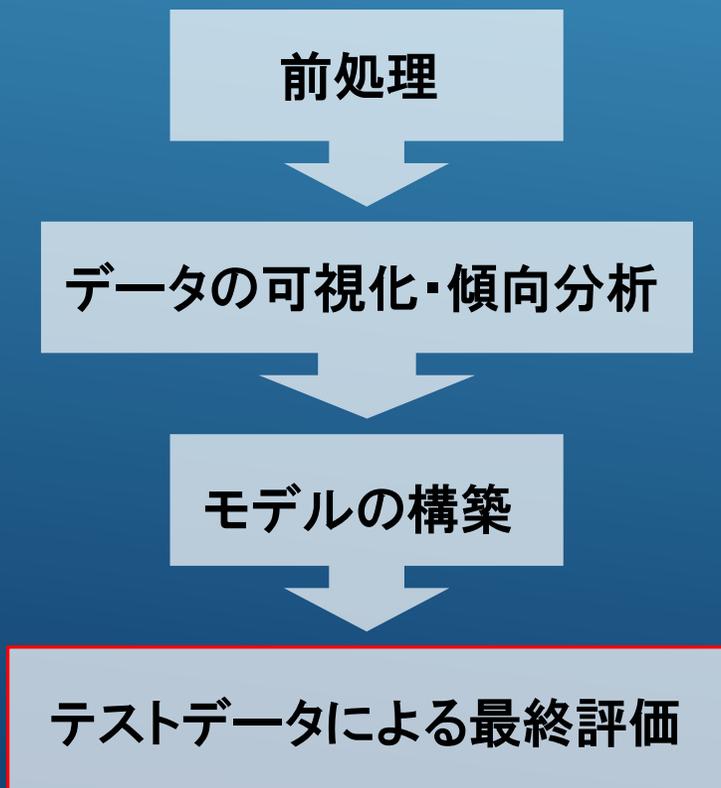
PBL (Project Based Learning)

The image features a solid blue background. In the bottom right corner, there are several white, parallel diagonal lines that create a sense of motion or a graphic element.

PBL(Project Based Learning)

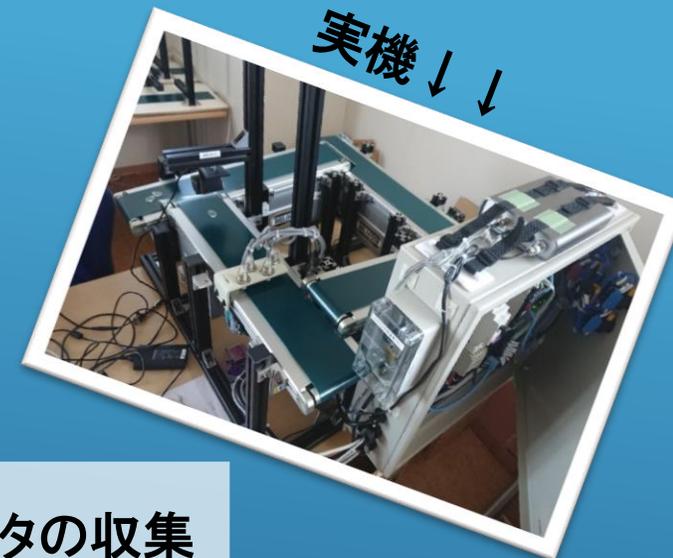
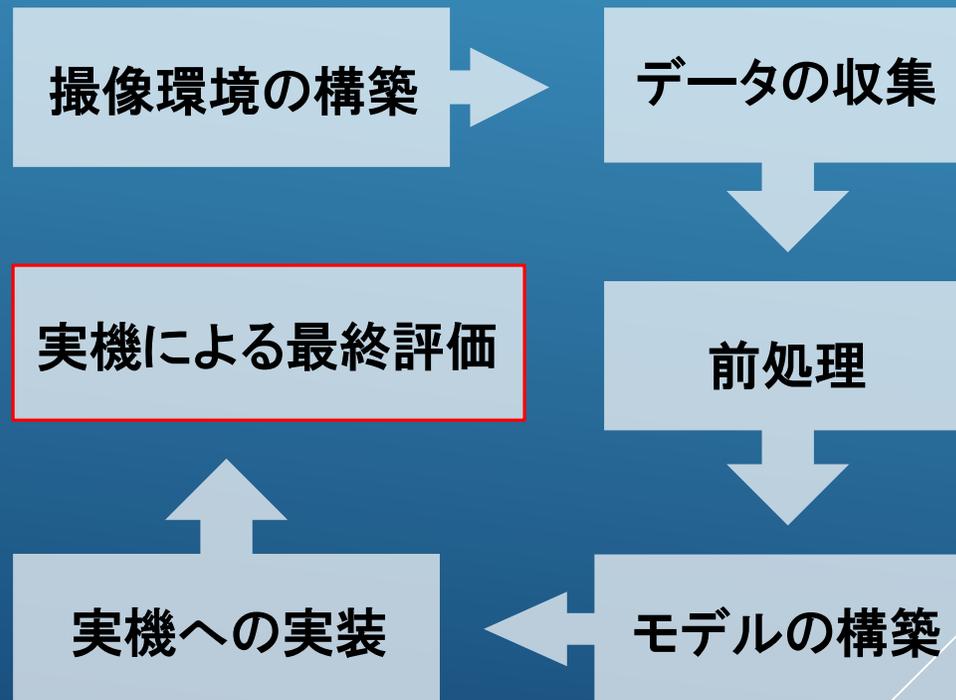
①時系列データ 5週間

《目的:異常の事前予測》

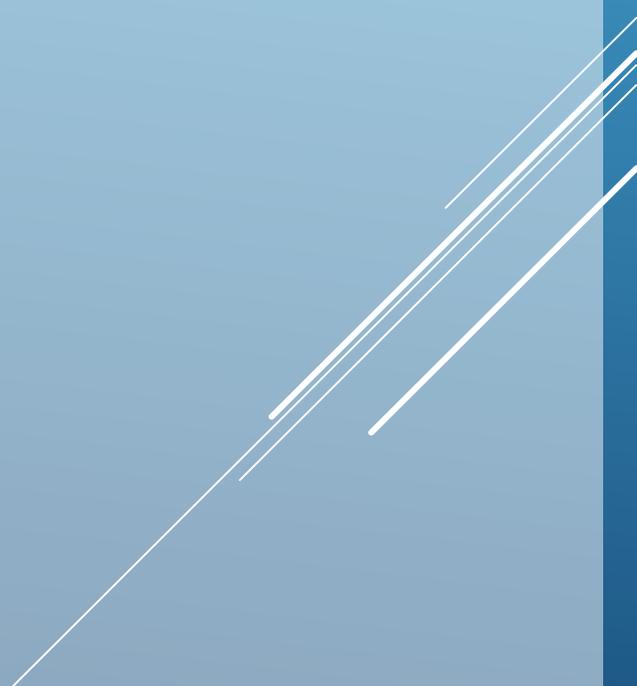


②画像データ 7週間

《目的:NG品の検出》



後半PBL PV動画



PBLを終えた受講生の声



・20代 男性

PBLを2種類今回実践しましたが、もっと回数を増やすことで進め方、対策法等をトライ&エラーで学んでいけると思いました
(ただ、時間がかかなり必要にはなっていますが…)
ただ、PBLを通して多くのことを学んだり、
チーム内の人とも仲良くなれたので有意義な時間でした

・30代 男性

チーム開発の難しさを経験できたのは非常に良かったです



・30代 男性

座学で学習したことよりも前の工程(今回で言えば撮影環境の構築や検証、アノテーション)についての試行錯誤、難しさの理解、考察をすることができた



受講生同士の交流



受講生同士の交流の重要性

受講生の声

・40代 男性
他社の困りごとを聞かせていただき、自社内での実用アイデアが広がった。



・40代 男性
グループディスカッションを通じての交流ができ、
機械学習を実務に取り込む問題点の気づきなどが持てた



・40代 男性
講義の内容を理解するための演習があり、自分でどこまで
理解できているか把握できて良かった。演習時に横の方と考え方を
相談する機会があり、色々な考え方があるなと気づきを持てた。



本講座で得られるのは知識やスキルだけではありません
志を同じくする仲間との縁は新たな気づきをもたらします

継続した交流促進への取り組み

テービーテックの役割

講座終了後も継続的に学びの場を提供し、
各社の横のつながりを得られる場を維持する

もくもく会の開催

本講座の受講生に限らず、プログラミングやAI・機械学習に興味のある方々に広く集まっていただき
自学習や情報交換・交流を行う会

学習教材として、Raspberry Pi 3・Dobot・
プログラミング制御が可能なドローンなどを用意
また、関連書籍をそろえて学習に打ち込める環境を提供します

<https://mokumokutoyota.connpass.com/event/151764/>

Slackによる交流推進

本講座ではコミュニケーションツールとして
Slackを使用しております
このチャンネルの一つとして豊田校1～3期・東京校
合同の交流チャンネルを作り、
期や会場を超えてAI・機械学習の最新情報を
共有したり質問をし合うなどの交流を促進しております

まとめ



まとめ

1.「製造業特化型データサイエンス集中コース」の位置付け

目的

年齢・経験の垣根を超え、製造業の現場で活躍できる
「現場」視点で考えることのできるデータサイエンティストを育成する

カリキュラム

製造業に必要な要素を凝縮し、演習・PBLによってより実践に近い環境下での学習を可能とする

2.継続した学習環境を提供することで製造業のデータサイエンス発展に貢献

・「製造業特化型データサイエンス集中コース」の継続

2020年5月に豊田校4期・東京校2期を開講予定

・受講終了後も継続した交流の推進

会場・期を超えた交流が出来るように定期的に集まる機会を設定し、気軽に集まれる場を提供する