

機械学習を用いた官能評価モデリングと車両運転性能の最適化

スズキ株式会社 岡島儀尚様

MathWorks Japan モハマッド ムキト

目次

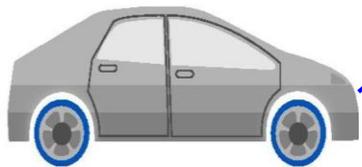
- 機械学習の取り組みにおける業界の動向 ✓
- **機械学習を用いた官能評価モデリングと車両運転性能の最適化**
(スズキ株式会社・岡島儀尚様による事例紹介)
- 機械学習を取り組むことに関する技術的なベストプラクティス

機械学習の時代を実感できるようになってきました

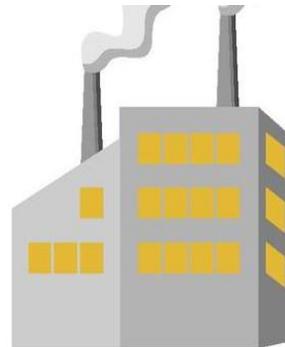
ランダム 官能評価 データ
フォレスト サイエンス
深層学習 SVM 交差検定 学習
機械学習 モデル
IoT ビッグデータ
ニューラルネットワーク 予知保全

機械学習は**予測**や**分類**のためだけに使うものでしょうか？

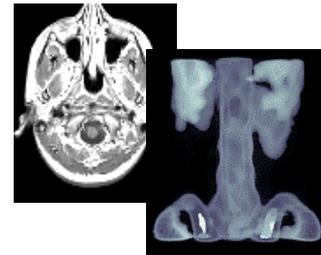
予測機能付き製品開発



障害物の
自動判断



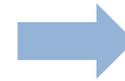
劣化度の
自動通知



健康状態の
自動判断

機械学習は予測や分類ではありません

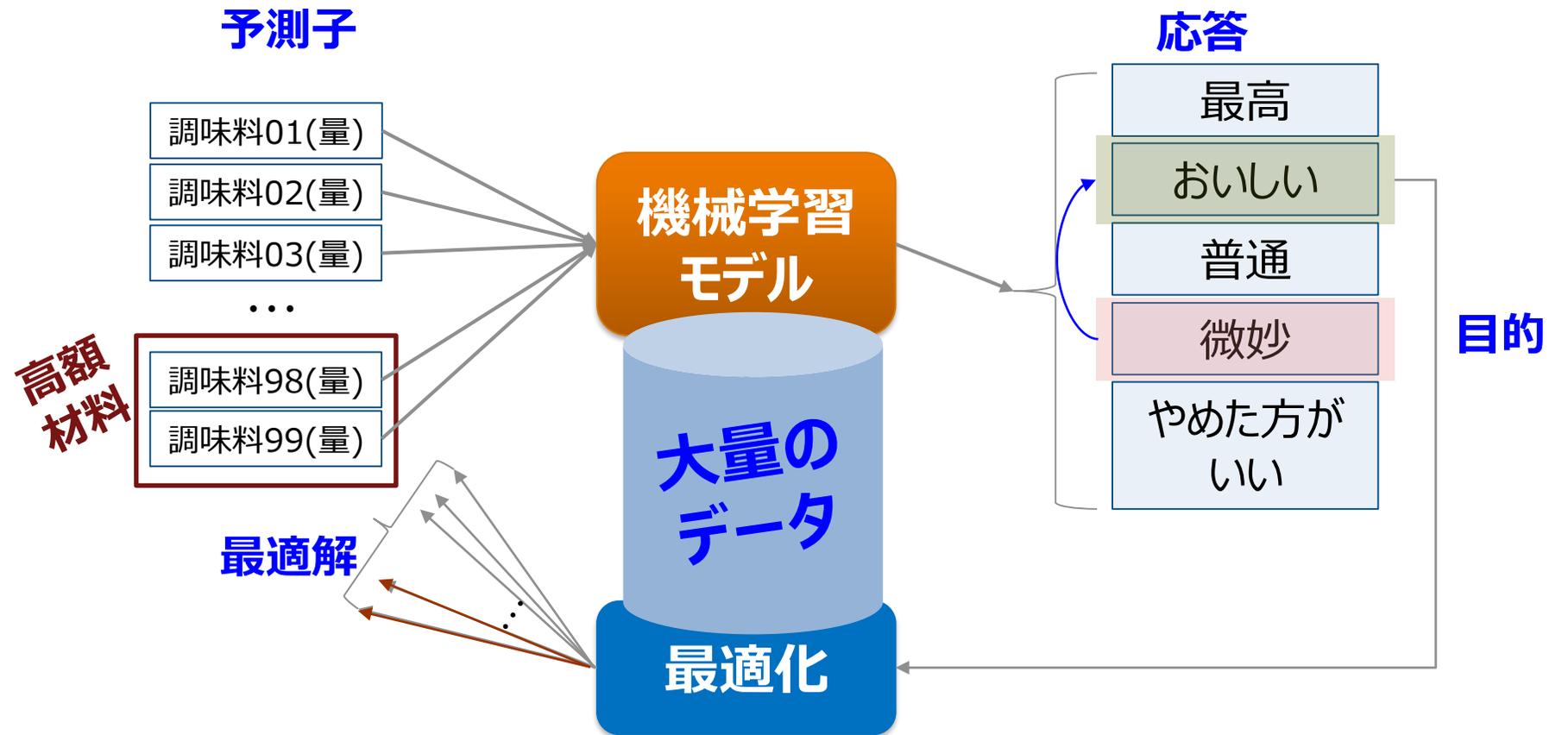
機械学習と最適化技術の組み合わせ



製品の一般的な性能を向上



自作カレー



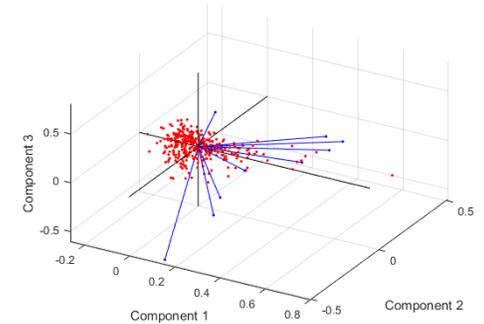
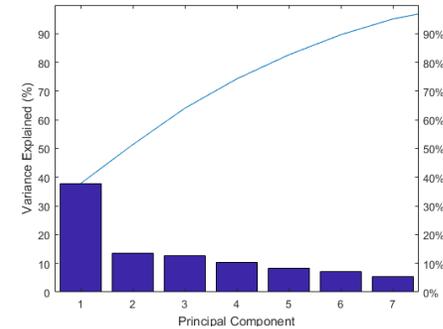
目次

- 機械学習の取り組みにおける業界の動向 ✓
- **機械学習を用いた官能評価モデリングと車両運転性能の最適化** ✓
(スズキ株式会社・岡島儀尚様による事例紹介)
- 機械学習に取り組むことに関する技術的なベストプラクティス ✓

データ前準備の各種項目について考えましょう

■ 次元削減・スケール調整

- 主成分分析 (PCA) 、回帰分析、正規化
- 寄与度評価 ⇒ 予測子最適化を簡易に
- 入力データの検査・除外

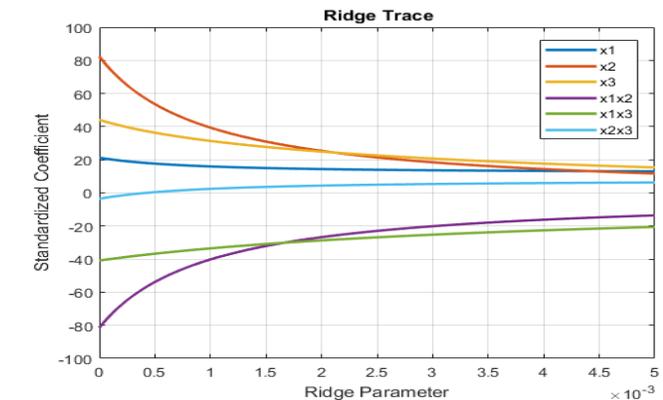


■ 前処理

- 異常値処理・欠測処理
- 複数のセンサーデータ間の同期
- 予測子に関する専門的な知識が大活躍する場
 - 取得データに普段どんなノイズが影響するか
 - 特定の処理でより効果的予測子を作ることが可能

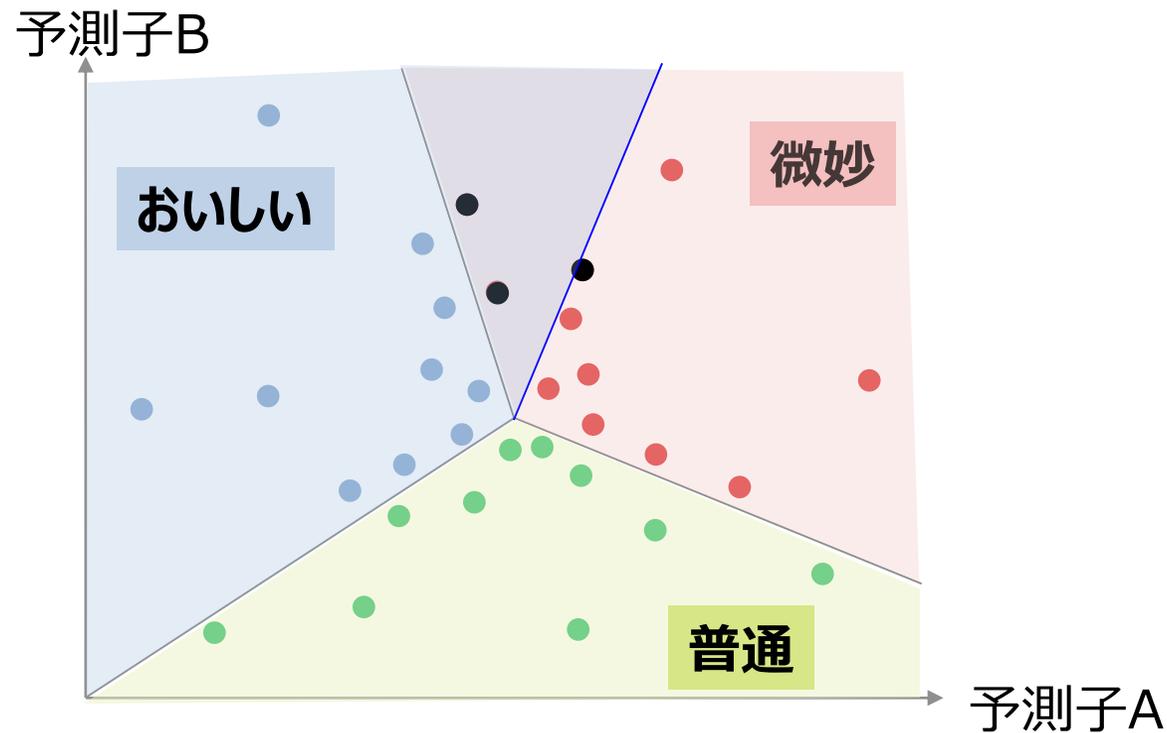
■ 学習の仕組み

- 用途に対して適切な方式
 - 分類? or 回帰?
 - 教師あり? or 教師なし?

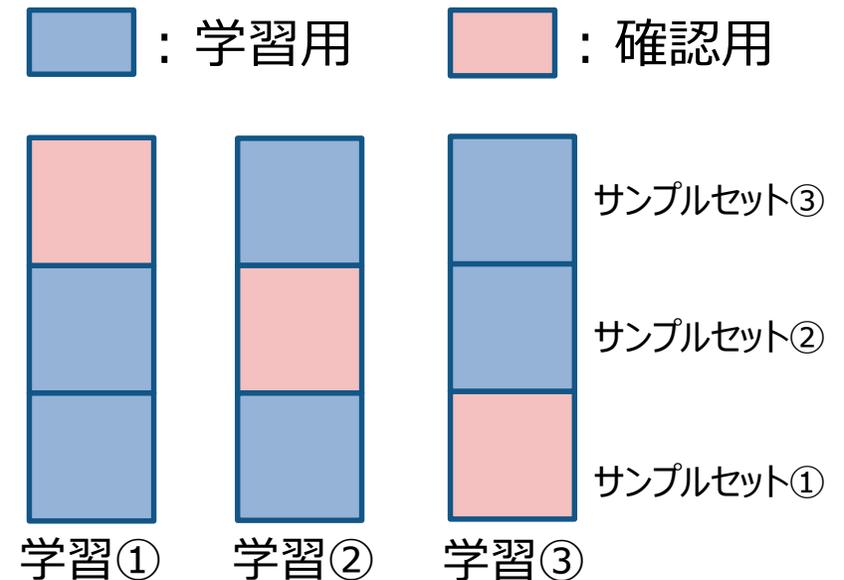


学習モデル評価には必ず交差検定をしましょう

- 本来の応答に関係を持つ特徴量が考慮されているかを明確化
- 交差検定しない場合のリスク
 - オーバーフィッティング：実際の予測精度より高い学習精度に見える
 - 応答と相関が弱い予測子が重要化される可能性も



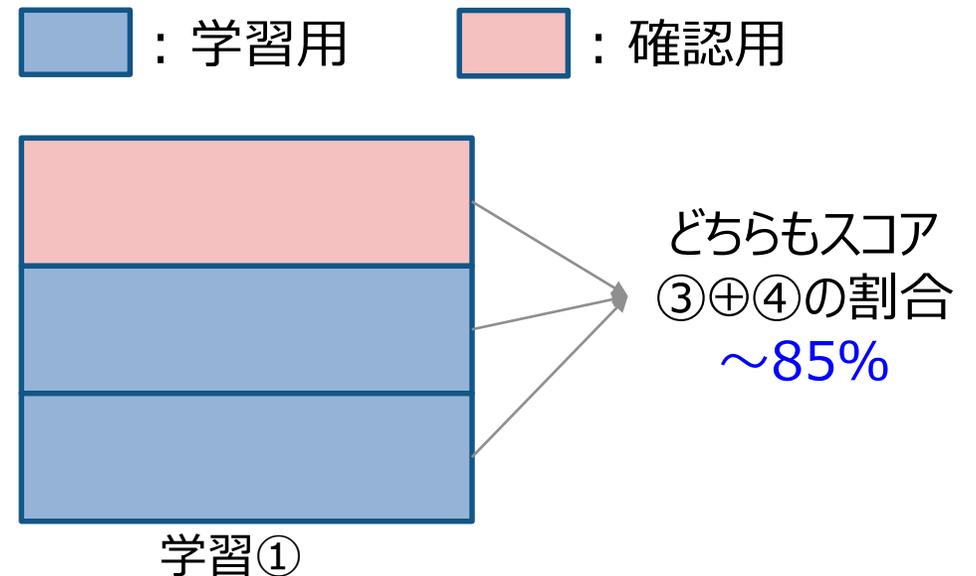
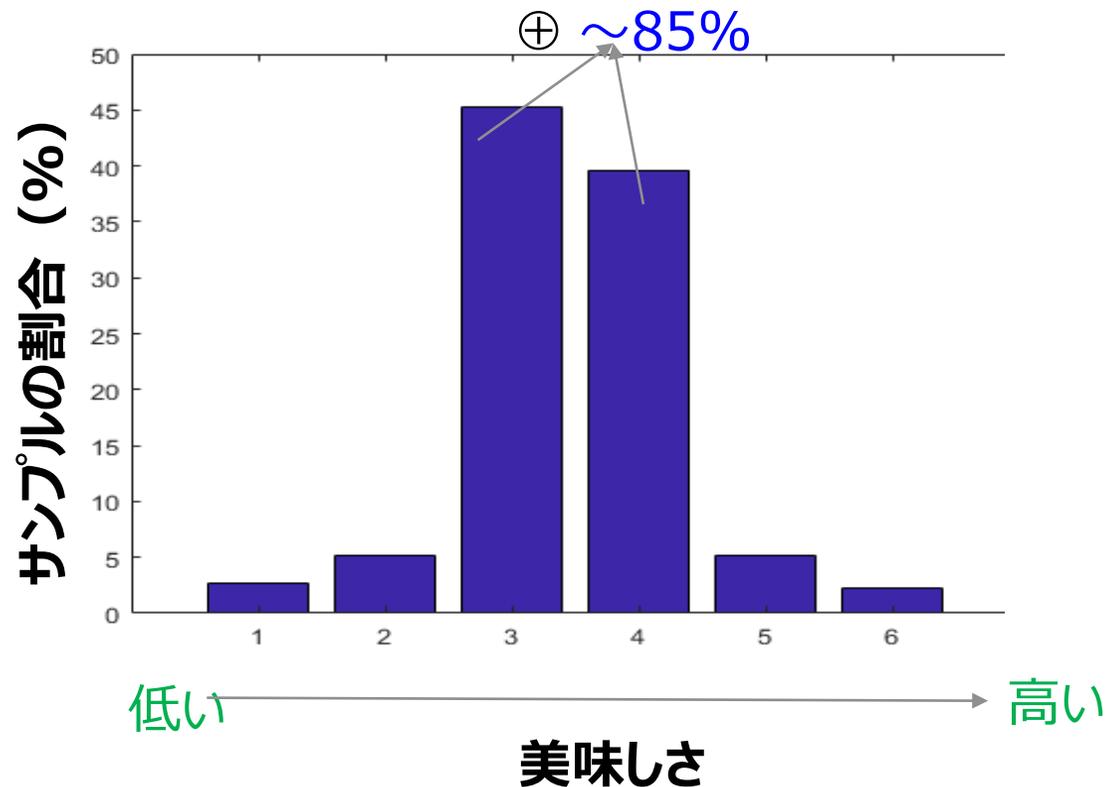
交差検定



学習用データの分布が応答スペクトルに一様になるようにしましょう

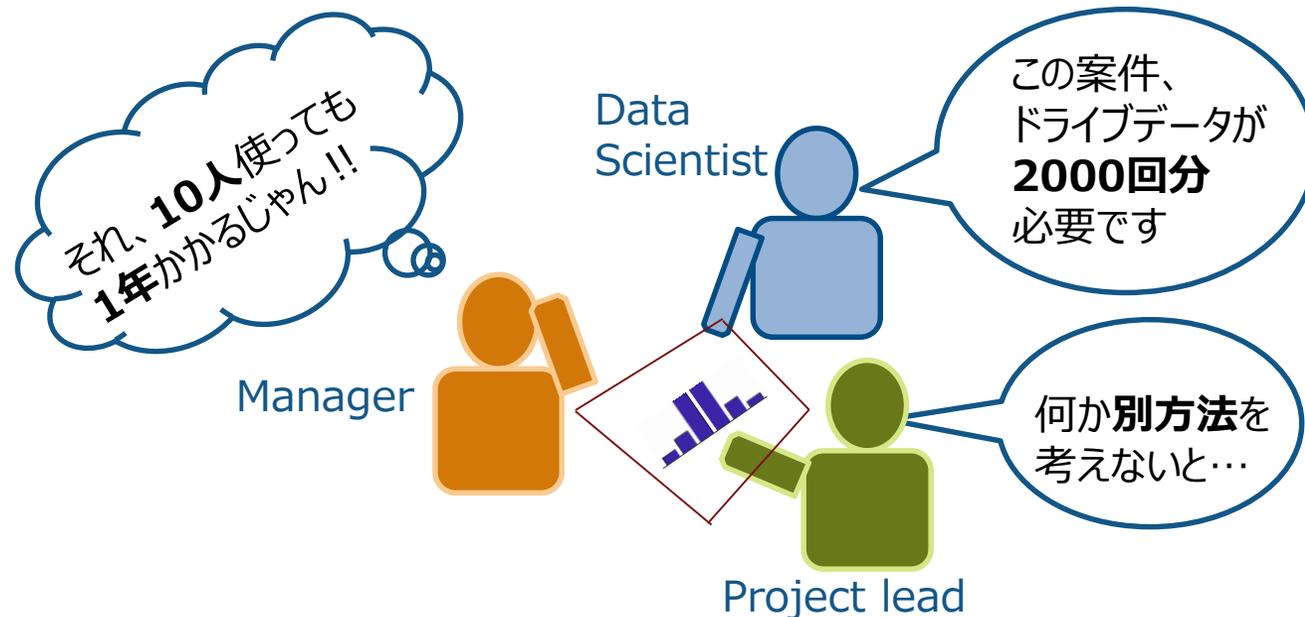
- 学習モデルが特定の応答に偏らない
- 交差検定の本来の目的を保つ

スコア③と④に偏った学習モデルでも
予測率が **~85%** になってしまう!!



データ取得についてシミュレーションを考慮しましょう

- データ取得仕組みのシミュレーションを適応
 - モデルベースデザイン (MBD) の適応
 - 人員、運営や時間のコストを大幅に削減
- 取得条件を変更してシミュレーションによる多数サンプル生成
 - 統計的なモデルによるノイズの追加でバリエーションの増加
 - 撮影条件 (カメラの角度・明るさ) を調整したシミュレーション

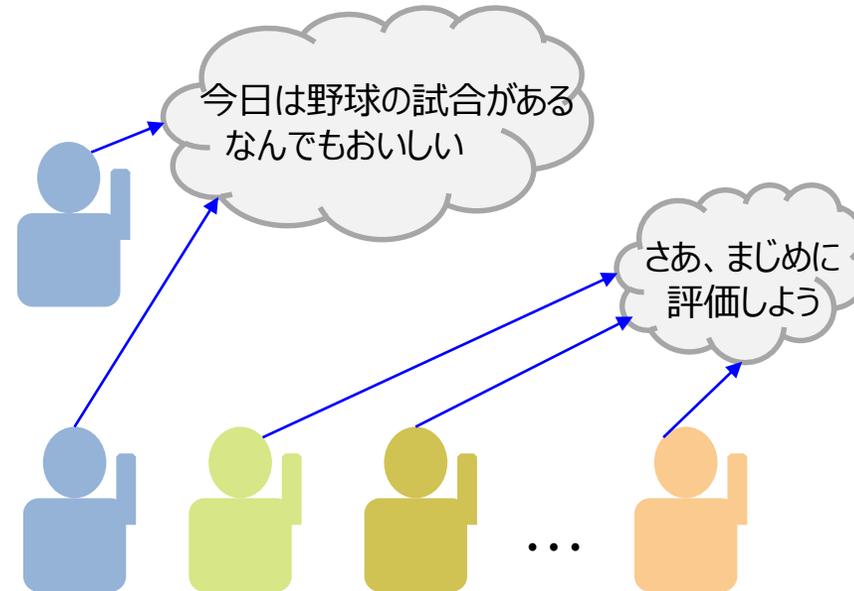


データラベリングに関するヒューマンエラーを抑える工夫をしましょう

- 「官能評価」では応答の定まった測定が難しい
 - 個人による偏りを避けることが非常に大事
 - 入力データを複数人がラベリング ⇒ ポーリング、平均値、など



一人試食



複数人試食

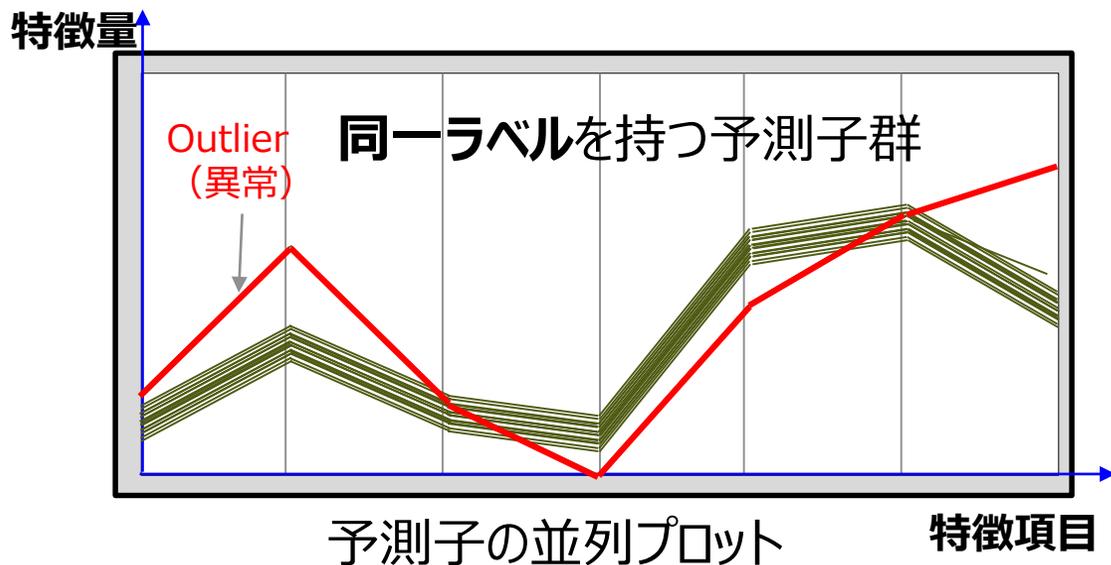
学習や評価にヒューマンエラーに対する対策を考慮しましょう

■ 予測子とラベルの「一貫性」をチェック

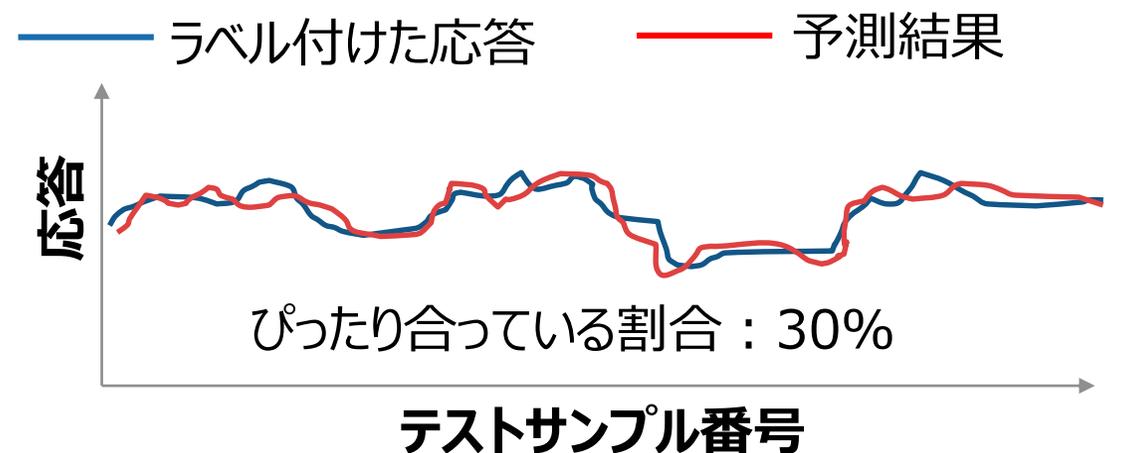
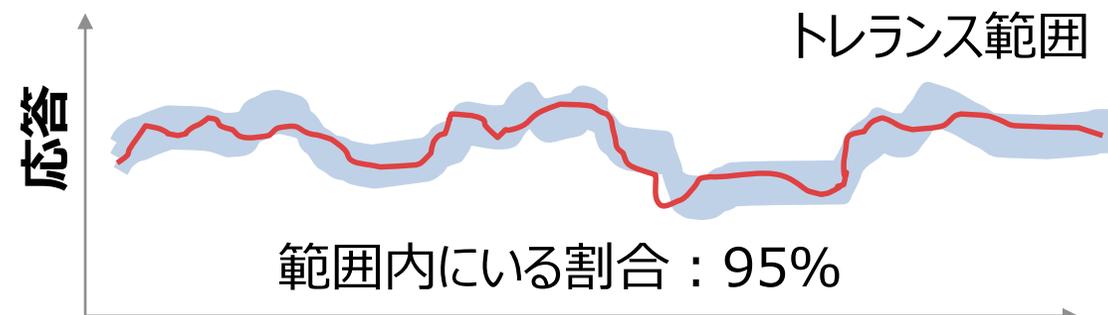
- ラベル修正、データ除外
- 同一ラベルの予測子カーブ間相関が十分?
- 高い相関の予測子同士のラベルが異なる?

■ 官能評価における「トレランス」概念の導入

- 人が「④」とスコア付けたが、実はスコアが「③」
- 予測検定に「許容範囲」導入でその合否を図る



トレランス導入によるモデル評価基準工夫



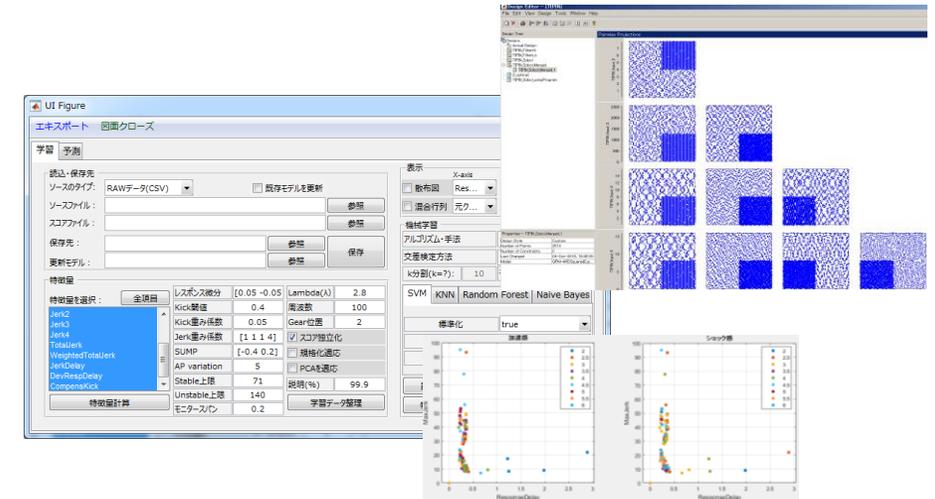
解析作業効率化や結果管理のためアプリケーション化をしましょう

- トライアンドエラーではそれぞれの項目について色々組み合わせをテスト
 - 【項目】：特徴量、前処理、学習アルゴリズム、評価方法、学習パラメータ、最適化、可視化方式

- 特徴量関連処理以外ほとんどは決まったワークフロー

- アプリケーション化に適切

- 開発負担が増えず、組み合わせ自由度が拡大化
- 短時間で複数の組み合わせの比較（可視化、データ表）
- 出力やレポートの管理が簡易化



- 複数の学習器を取り組んだシステム化

岡島様のお話



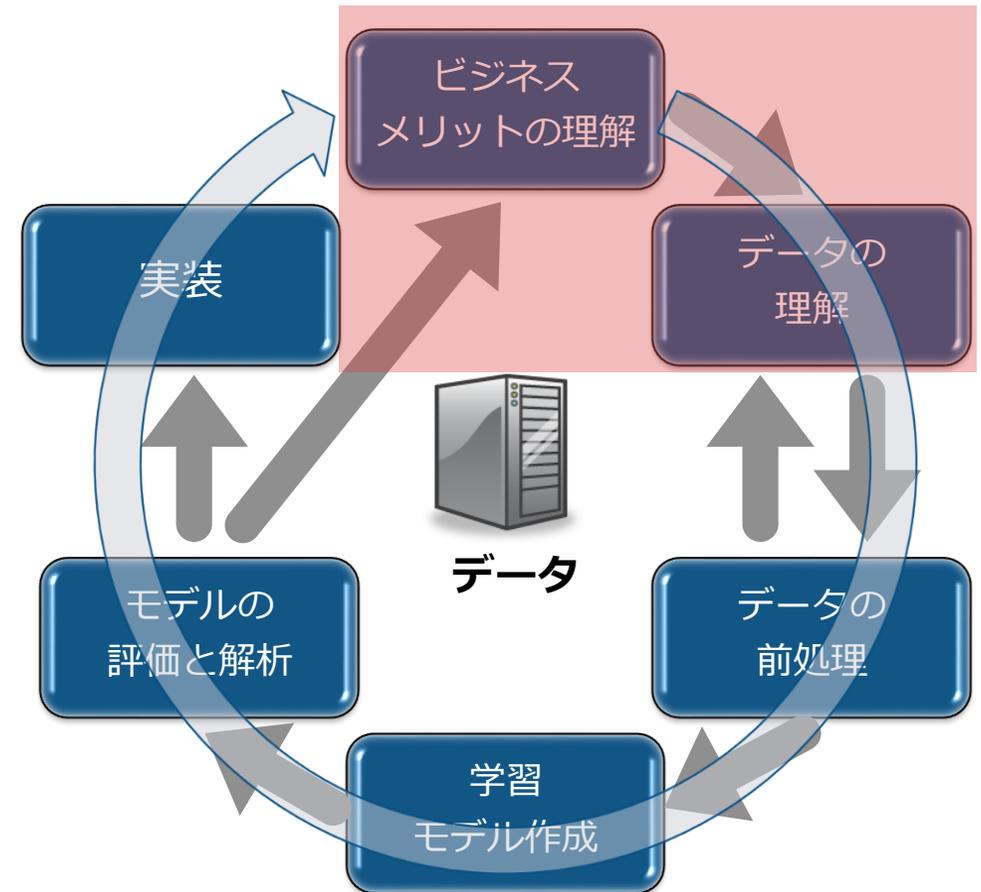
データ解析フローを活用しましょう

■ CRISP-DM (Cross-Industry Standard Process for Data Mining)

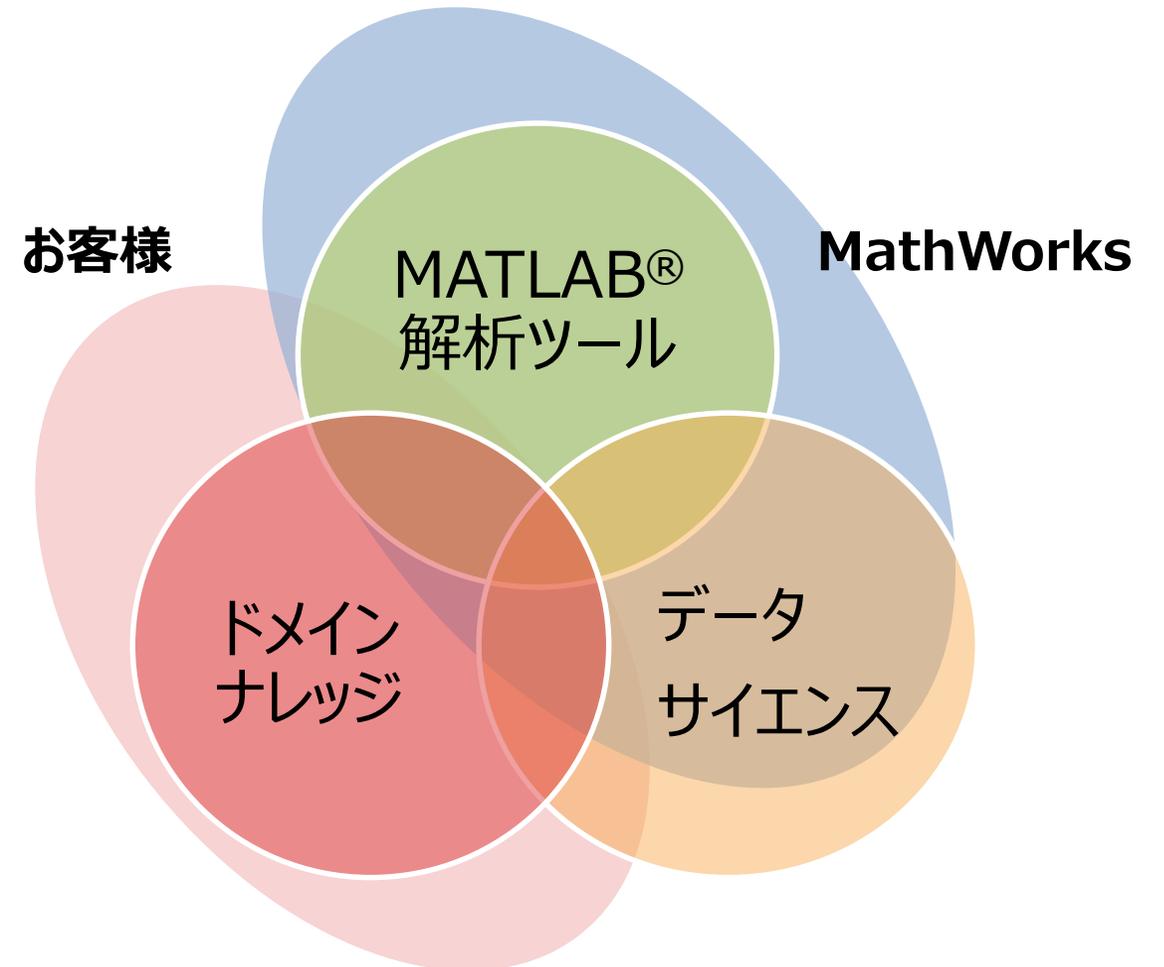
- データマイニングコンソーシアムで開発された方法論
- 繰り返し実行によって目標達成へ
 - トライアンドエラー
 - PDCA (Plan ⇒ Do ⇒ Check ⇒ Act)

■ データを理解することが重要

- どんなデータを持っているか
- 可視化による特徴把握
- 専門知識を活用 (データサイエンスとのSynergy)
- 予測結果や最適解のメリット解析



MathWorks Japanコンサルティングは機械学習の一貫的なサポートを提供



是非、機械学習の取り組みにMATLABを活用しましょう

- 予測機能付き製品開発とともに、多くの業界では機械学習の最適化の組み合わせによるあらゆる製品の一般的な性能向上
- MATLABプロダクトファミリーが「**機械学習**」、「**モデルベース開発(MBD)**」、「**最適化**」や「**画像・信号処理**」のため強力なプラットフォームを提供
- MathWorks Japan エンジニアリングサービスのサポートによって目的達成をスピードアップ
- 是非、MathWorks Japan のウェブページのリソースをご活用ください
<http://www.mathworks.com/>