

Section 5

- 機械学習エンジニア育成のために

参考事例

- データサイエンスアワード 2017 最優秀賞
 - 「データサイエンスと働き方改革
～安全で効率的なメンテナンスを目指して～」
 - 学校法人産業能率大学総合研究所
 - 東京地下鉄株式会社
- http://www.datascientist.or.jp/symp/2017/pdf/award_tokyometro.pdf

問題点

- 地下鉄の安全な運行のため、日々の膨大な点検
 - 点検結果の集約に時間がかかる（最大3ヶ月）
 - トンネル補修計画立案の難しさ
 - 優先的に補修すべき箇所を見つけたい
 - 目視による点検の不確実性
 - 異常な状態の見逃しをなくしたい

解決法

- 土木構造物検査データ管理システムの構築
 - 検査工程管理と検査結果の迅速な共有
- 機械学習の業務への組み込み
 - ベイズ推定による確率モデルを用いたリスク評価
 - ベイジアンネットワークによる異常見逃しの防止
- データサイエンティスト育成のための研修
 - 様々な部署からメンバーを選抜し、 3 年間実施

機械学習技術者に求められるもの

- インフラ
 - データの取得、管理、計算資源の確保ができる
- 実践
 - 対象業務を深く理解し、収集するデータの質を高めることができる
 - どの程度の性能で、何ができるかを見極めることができる
- 理論
 - 性能勝負の段階で、いかに多くの引き出しを持っているか

機械学習技術者に求められるもの

- インフラ・実践・理論のすべてを完全にカバーできる人はいない
- 上記3つのいずれかに軸足を置き、他の2つは何をやっているかがわかる程度になっておく
- 技術も変化するので、継続的な勉強が必要

機械学習技術者を育成するには

- 事業のコア部分に機械学習を取り込む
 - 長期的視点で育成
- 勉強時間を確保する
 - 働き方改革が始動した今がチャンス
- よいコミュニティの形成を促す
 - 社内：何を秘密にするかルールを明確に
 - 社外：ロジスティックスサポートを惜しまずに

本日のまとめ（ 1/4 ）

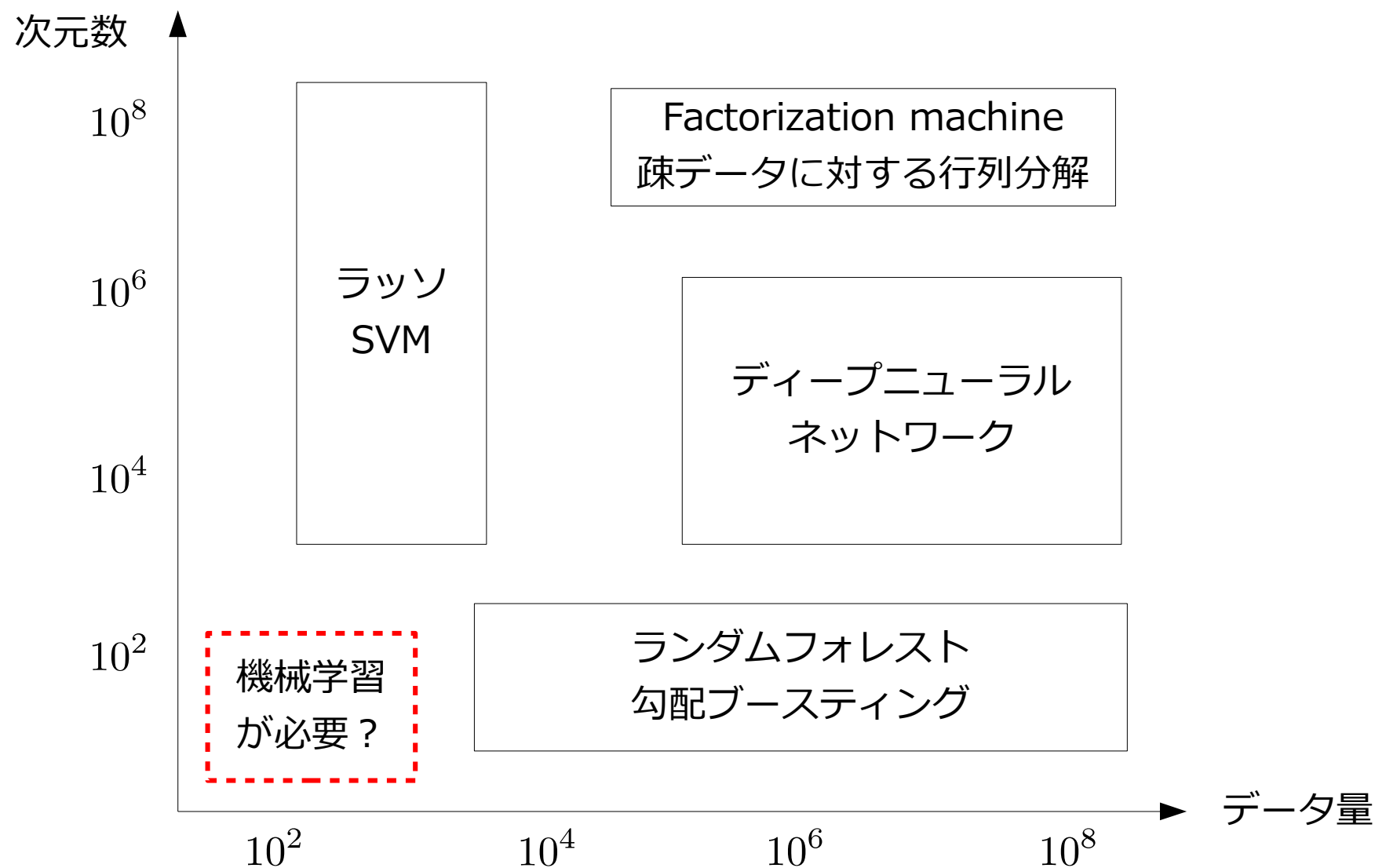
- 人工知能
 - 人間の知的作業の一部を置き換える技術
- 機械学習
 - 人工知能を実現するための技術のひとつ
 - データがなければ始まらない
- 深層学習
 - 特定の問題に対して、高い性能を実現する機械学習技術のひとつ
 - 少ないデータでもなんとかなる可能性が見えてきた

本日のまとめ（ 2/4 ）

- 機械学習でできること
 - 識別：入力を予め決めたクラスに分類する
 - 決定木、ロジスティック識別、SVM、勾配ブースティング、ディープニューラルネットワーク
 - 回帰：様々な特徴からターゲットの値を予測する
 - リッジ回帰、ラッソ回帰、モデル木
 - モデル推定：教師なしデータからのモデルの作成
 - クラスタリング、異常検知
 - パターンマイニング：有用な規則の抽出
 - アプリアリ、行列分解

本日のまとめ（3/4）

- 問題を正確に定義し、適用する手法を見極める



本日のまとめ（ 4/4 ）

- 機械学習技術者の育成
 - コア事業が機械学習なしでは成り立たないようにする
 - ほとんどの事業がそうなるべきでは？ ...
- 現実的な教育計画
 - 長期に渡った十分な投資を
 - 月 1 回 2 時間 × 半年で育成？
 - 基本的な数学は必要だが、範囲はそれほど広くない
 - コードのみで学んでゆくことも可能
 - よきパートナーとなる大学を見つけてください