機械学習の活用に向けて

統計数理研究所
松井知子

機械学習

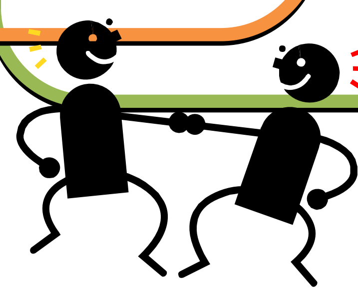
- データから知識を抽出する方法論〜データからの帰納
- McKinsey Global Instituteのレポート(2011)
  - 機械学習の重要性を言及
    「機械学習はイノベーションの次の大きな波を呼ぶ」
- P. Domingos, “A few useful things to know about machine learning,” Comm. of The ACM, 55(10), 2012
  - 機械学習の活用に必要な知識を紹介

機械学習

- Learning = Representation + Evaluation + Optimization
  [Domingos, 2012]
- 機械学習～
  $\underbrace{\text{コーディング＋モデリング}}_{\text{Representation}}$＋目的関数＋最適化

アプリケーションドメインの専門家

機械学習の専門家



機械学習

- モデリング
  - 線形回帰/識別モデル、ニューラルネットワーク、SVM、ロジスティック回帰、グラフィカルモデル、混合ガウス分布、隠れマルコフモデル、決定木 など
- 目的関数
  - 最小二乗誤差、最小KLダイバージェンス、最大マージン、最尤、最大事後確率 など
- 最適化
  - 連続最適化：最急降下法、共役勾配法、ニュートン法、EMアルゴリズム など
  - 離散最適化：ビーム探索、greedy探索 など

機械学習

- David Hume（1711-1776、哲学者）

  「知識の起源を知覚により得られる観念にあるとすれば、確実な知に人間本性が達することは原理的に保証されない」

  ➡何らかの知識や仮定の必要性

- No-free-lunch定理 [Wolpert and Macready、1995]

  「どの学習機械に対しても、ランダム推測の方が優れている目的関数が少なくとも一つ存在する」

ITSのデータ

- プローブ情報、カーナビ情報、車両情報、持ち込み端末のセンサ情報など
- ビッグデータの三つの「V」
  - Volume（膨大なデータ）
  - Velocity（高速度でのデータ入出力）
  - Variety（多種多様なデータ）

カーネル法

カーネル法

- カーネル関数（データ同士の類似の度合いをはかる関数）を利用して識別や回帰の問題に対処する。
  - データを高次元空間に写像した上で処理する。

カーネル関数〜次元の呪いの克服

- カーネル関数は多種多様なデータに対して設計できる。
  - 連続値／離散値
  - ベクトル、系列、グラフ、木、…
- カーネル法の例
  - サポートベクターマシン、カーネルロジスティック回帰マシン、ニューラルネット、カーネルPCAなど

花塚泰史（（株）ブリヂストン/総研大）、樋口知之、松井知子（統数研）

2012 IET International Conference on Information Science and Control Engineering (ICISCE 2012)

DEVELOPMENT OF THE ROAD-SURFACE CONDITION CLASSIFICATION SYSTEM

カーネル法による路面状態の判別

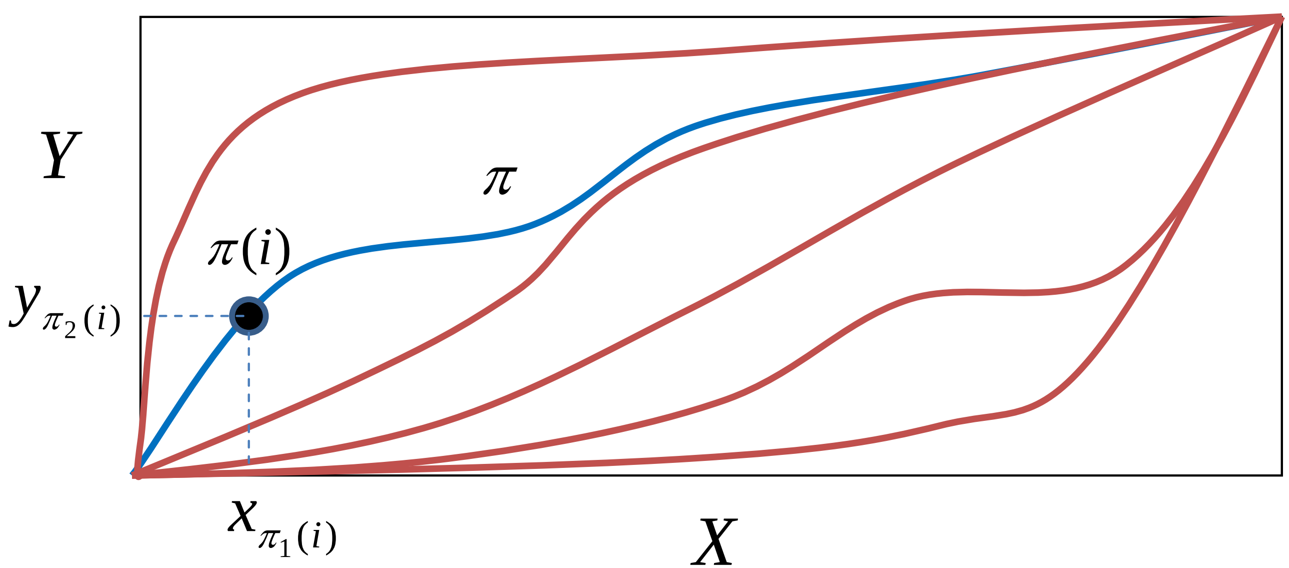
- 入力：タイヤ振動の時系列信号
  （ケプストラムベクトル時系列）
- 出力：路面状態（ドライ／ウェット／圧雪／凍結）
- 判別法：
  - サポートベクターマシン（SVM）＋グローバルアライメント（GA）カーネル
  - 隠れマルコフモデル（HMM）

GAカーネル

[Cuturi, Vert, Birkenes and Matsui, ICASSP2007]

- 可変長、観測値のばらつきを考慮
- 全パスπに沿って累積距離を計算

$$S(\pi) = \sum_{i=1}^{|\pi|} \phi(x_{\pi_1(i)}, y_{\pi_2(i)}) \qquad \phi(x,y) = -\|x-y\|^2$$



GAカーネル

- 定義：
$$K_{GA}(\mathbf{x},\mathbf{y}) = \sum_{\pi \in A(\mathbf{x},\mathbf{y})} \exp\left(S(\pi)\right)$$
$$= \sum_{\pi \in A(\mathbf{x},\mathbf{y})} \exp\left(\sum_{i=1}^{|\pi|} \phi(x_{\pi_1(i)}, y_{\pi_2(i)})\right)$$
$$= \sum_{\pi \in A(\mathbf{x},\mathbf{y})} \prod_{i=1}^{|\pi|} k(x_{\pi_1(i)}, y_{\pi_2(i)})$$

[$k = e^{\varphi}$]

実験

- 実験条件
  - 利用したタイヤ：4種類（うち1種類を学習、残り3種類をテスト）
  - コーディング：8次元ケプストラム＋Δ＋ΔΔ　〜24次元ベクトル
    （窓長3ms、フレームシフト2ms）
- 結果
  - GAカーネルによる方法：　97.2%
  - HMMによる方法（4状態）：　85.1%

まとめ

- 機械学習はITSを含め、ビッグデータ解析の重要なツール:
  - うまく活用するためには、対象ドメイン／アプリケーションに即した知識や仮定が必要
  - データの多様性に対応できる方法、カーネル法を紹介
  - 「カーネル法による路面状態の判別」の研究例を紹介