

Item-multisetカーネルに対するスパースなランダム特徴

カーネル法: 超高次元(時には無限次元の)空間上でのデータ分析を,
次元空間における座標を陽に計算せずに可能にする方法

超高

- ☺高次元空間での内積(カーネル関数)さえ計算できれば良い
- ☹データ数に対してスケールしない(データ数が多いと動かせない)

ランダム特徴: カーネルに対応する高次元空間を近似するそこそこの次元の空間
にランダムに飛ばし, その空間の上で線形モデルを直接動かす

本研究の貢献

Item-multisetカーネルに対するスパースなランダム特徴を提案

- Item-multisetカーネル: Itemsetカーネルや内積カーネルを含む広いカーネル
- 写像に用いる行列がスパース→写像の計算が高速・省メモリ化
- 写像後のランダム特徴行列がスパース→写像後の学習が高速化
- 大規模スパースデータに対してもランダム特徴が適用可能に