

02

複数の問題に共通して重要な情報の組合せを発見

～共通因子を効率的に学習する低ランク回帰技術:MOFM～



どんな研究

今まで一つの目的変数の予測に限られていた「Factorization Machines」を**複数の目的変数の予測**に拡張した研究です。提案法はデータを表す特徴から**各目的変数に対して有用な組合せ**を見つけて高精度なモデルを作ります。

どこが凄い

提案法は、各目的変数に**共通の基底**を用いて特徴の組合せを表現し、重みを求めるため、次元数が膨大になっても効率的に学習できます。また、初期値に依存せず**大域的最適解**を得られることが理論的に保証されています。

めざす未来

応用先として、医療診断、推薦システム、遺伝子解析などに利用可能です。今後は、**理論的性質**（表現力、汎化能力など）に関する理解を深め、提案法を再帰的に適用することで**深層学習**での効率的な計算に活用していきます。

Multi-Output Factorization Machines (MOFM): 複数タスクのモデルを同時に学習し重要な特徴の組合せを発見

- 汎用的(様々な予測問題に適用可能)
- タスク間の相関関係を活用(各重み行列を共通の基底で低ランクに分解)
- モデルのパラメータ学習が容易(解が初期値に依存せず一意に求まる)
- 膨大な数の組合せを扱える(高次元データに適用可能)

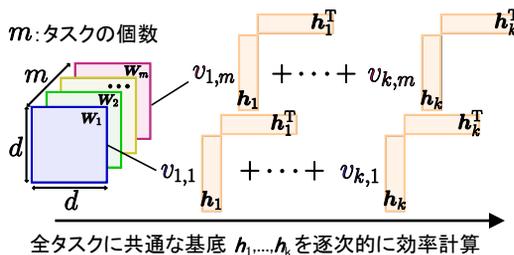
$$\hat{y}_m = \mathbf{x}^T \mathbf{w}_m + \mathbf{x}^T \mathbf{W}_m \mathbf{x}$$

↑ m番目のタスク

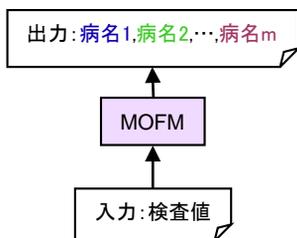
$$\hat{y}_2 = \mathbf{x}^T \mathbf{w}_2 + \mathbf{x}^T \mathbf{W}_2 \mathbf{x}$$

$$\hat{y}_1 = \mathbf{x}^T \mathbf{w}_1 + \mathbf{x}^T \mathbf{W}_1 \mathbf{x}$$

膨大!

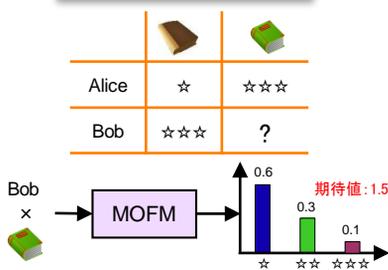


応用例1: 医療診断



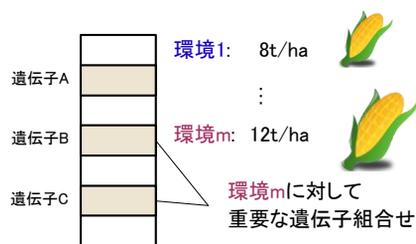
複数の病名(1... m)を診断し共通な要因(検査値の組合せ)を発見

応用例2: 商品推薦



評価点(1... m)に対して確率を予測し期待値を用いて書籍推薦

応用例3: 収穫量予測



各環境(1... m)に対するトウモロコシの収穫量を遺伝子の組合せから予測

関連文献

[1] M. Blondel, A. Fujino, N. Ueda, "Convex Factorization Machines," in *Proc. European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 2015.

[2] M. Blondel, V. Niculae, T. Otsuka, N. Ueda, "Multi-output Polynomial Networks and Factorization Machines," in *Proc. Neural Information Processing Systems*, 2017.

担当者

ブロンデル マチュー (Mathieu Blondel) 上田特別研究室