

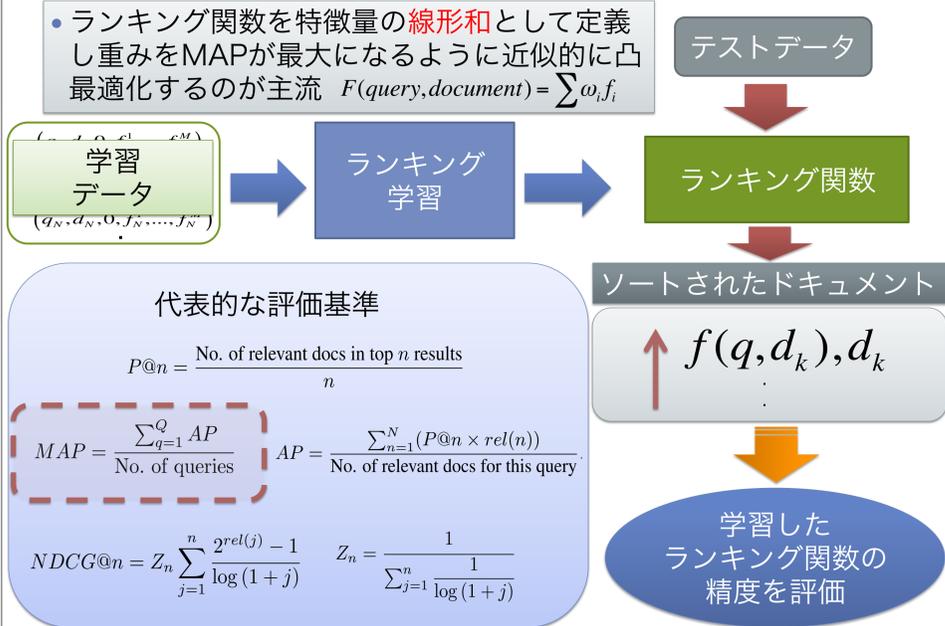
確率モデルGPによるランキング学習の提案

佐藤浩之, ダヌシカボレガラ, 伊庭斉志 (東京大学)

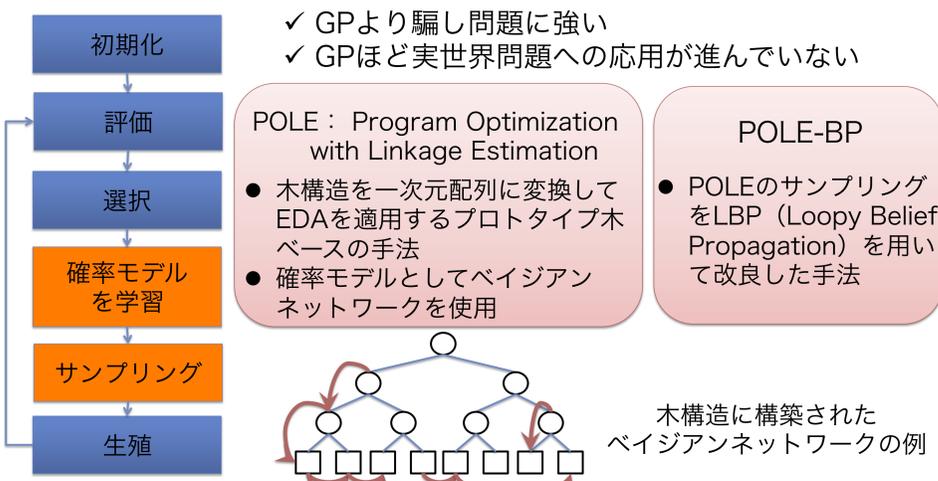
研究の背景

- 検索エンジン
 - 検索語 (クエリ) と関連のあるウェブページ (ドキュメント) が検索結果の上位にきてほしい
 - そこでランキング関数の出力でウェブページをソート
- ランキング関数
 - $f(query, document) \Rightarrow$ 関連度 (実数値)
- ランキング学習
 - ランキング関数を教師あり学習すること
 - ユーザの満足度や検索エンジン運営企業の収益に関わるため重要

ランキング学習の流れ



確率モデルGP

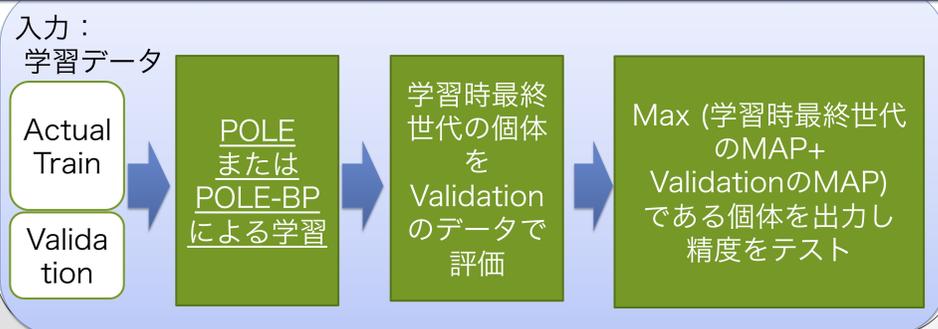


提案手法

確率モデルGPによるランキング学習

メリット	詳細
<ul style="list-style-type: none"> 直接MAPを最適化できる ランキング学習には凸な損失関数が存在しないため (C. Calauzènes et al. 2012) 非凸最適化が有効 表現能力が高い非線形な関数を学習できる 確率モデルGPの実世界問題への応用の促進 	<ul style="list-style-type: none"> 個体: ランキング関数 終端ノード <ul style="list-style-type: none"> S_v: 特徴量に対応する変数 S_c: 定数値 関数ノード <ul style="list-style-type: none"> S_f: {+, -, ×} 適合度: MAP

RankPOLE (POLEによる学習), RankPOLE-BP (POLE-BPによる学習)



LETOR (version 2.0, TD2003)



実験と考察

- ✓ 個体数5000, 世代数50, 木の深さ8, 結果は3試行の平均
- ✓ 確率モデルGPではノードの種類が多いと確率モデルの学習がうまくいかないため, 特徴量 (変数ノード) の削減を検討した

条件1: 44種類全ての特徴量を使用 (条件1のみ定数ノードは{0.1, 0.2, ..., 1.0}の10種類, 条件2~4では定数ノードは{0.2, 0.4, ..., 1.0}の5種類)

Depth	RankPOLE				RankPOLE-BP			
	5	6	7	8	5	6	7	8
MAP	0.223	0.227	0.222	0.209	0.235	0.245	0.236	0.226

条件2: DEで線形和を最適化する手法 (RankDE, Danushka et al.) で重みの絶対値が2以上となった11種類の特徴量のみを使用 (特徴量ID: 5, 7, 8, 9, 10, 11, 12, 18, 21, 23, 39)

Depth	RankPOLE				RankPOLE-BP			
	5	6	7	8	5	6	7	8
MAP	0.247	0.242	0.252	0.256	0.251	0.258	0.255	0.263

条件3: GP (交叉率0.8, 突然変異率0.1, 個体数600) で全ての特徴量を使いランキング学習を40回行い, 出力された個体において出現回数の上位10種類の特徴量のみを使用 (特徴量ID: 7, 8, 13, 17, 24, 27, 34, 37, 39, 42)

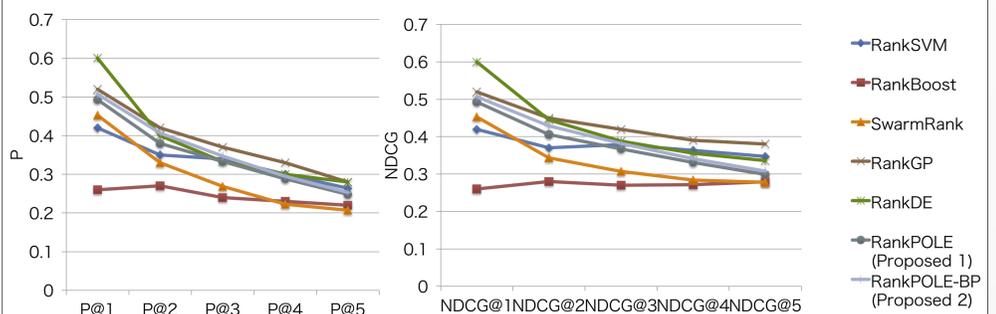
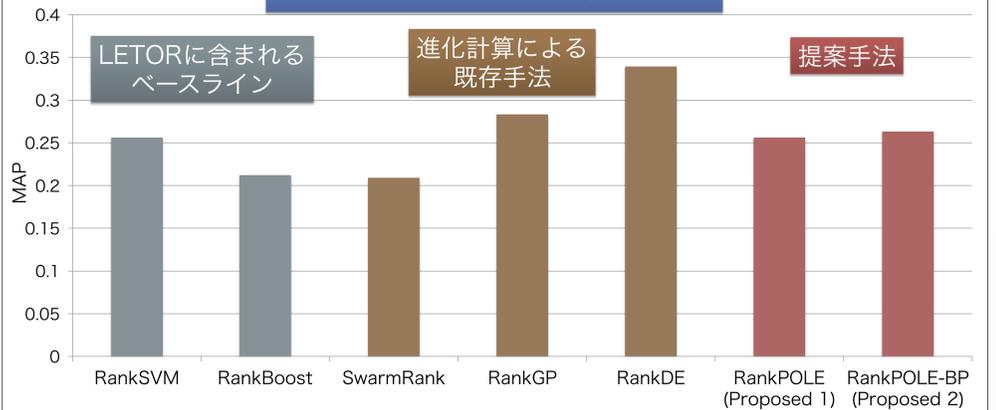
Depth	RankPOLE				RankPOLE-BP			
	5	6	7	8	5	6	7	8
MAP	-	0.231	-	0.248	-	0.236	-	0.244

条件4: 条件3と同様にして, 出現回数とその特徴量が含まれる個体のMAPの両方を考慮し選出した10種類の特徴量のみを使用 (特徴量ID: 7, 8, 11, 13, 17, 24, 27, 34, 39, 42)

Depth	RankPOLE				RankPOLE-BP			
	5	6	7	8	5	6	7	8
MAP	-	0.224	-	0.239	-	0.229	-	0.236

- ◆ ノードの種類の削減は有効に働き, 性能の向上をもたらしている
- ◆ しかし, 線形和を最適化する手法の知見を活用して特徴量を削減した条件2より, 非線形関数を最適化する手法の知見を活用した条件3, 4のほうが性能が低い
- ◆ GPによる特徴量選択が上手く働いていないのは突然変異率を固定しているから? (GPによる既存手法 (Jen yuan Yeh et al.) が存在するが, 突然変異率を動的に変化させている)

既存手法と提案手法の性能比較



まとめと今後の課題

- ✓ 確率モデルGPをランキング学習に応用したところ, ベースラインやPSOを用いた既存手法を上回る性能を示した
- ✓ 確率モデルGPをシンボルの種類が多い実世界問題へ応用する際には, シンボルの種類の削減が有効である
- ✓ 動的に突然変異率を変化させるRankGP (Jen yuan Yeh et al.) による特徴量の選択を行いたい
- ✓ 確率モデルGPとGPのマイグレーションによるランキング学習を行いたい

参考文献

- Danushka Bollegala, Nasimul Noman, and Hitoshi Iba. RankDE: learning a ranking function for information retrieval using differential evolution. In Proceedings of the 13th annual conference on Genetic and evolutionary computation, GECCO '11, pp. 1771-1778, New York, NY, USA, 2011. ACM.
- Clément Calauzènes, Nicolas Usunier, and Patrick Gallinari. "on the (non-)existence of convex, calibrated surrogate losses for ranking". In P. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou, and K.Q. Weinberger, editors, Advances in Neural Information Processing Systems 25, pp. 197-205, 2012.
- Jen yuan Yeh, Jung yi Lin, Hao ren Ke, and Wei pang Yang. Learning to rank for information retrieval using genetic programming, 2007.