

Web調べ学習における学習者の背景知識に基づいたランキングの最適化

情報科学科 加藤 敦樹

指導教員：小林 邦和

1 はじめに

近年、情報検索技術の普及により、あらゆる場面で Web 検索が利用されている。特に、大学生には Web 検索を利用した調べ学習が浸透している。Web 検索を利用した調べ学習では、サイトごとに説明の仕方は多種多様であり、その中から学習者にとって有益なサイトを探し出すことは難しい。

そこで本研究では、学習者の背景知識に基づいて検索ランキングを最適化することによって、学習者に適したランキングで検索結果を提示し、Web 調べ学習での学習効率の向上を目指す。

2 提案手法

本研究では、学習者が過去に閲覧した Web サイト内のトピックとその Web サイトに対する評価から、学習者のトピックごとの習熟度をモデル化する。そして、そのトピックごとの習熟度に基づく理解の可能性の高い Web サイトの順で Google の検索ランキング上位を入れ替えて学習者に提示するシステムを提案する。提案システムの構成を図 1 に示す。

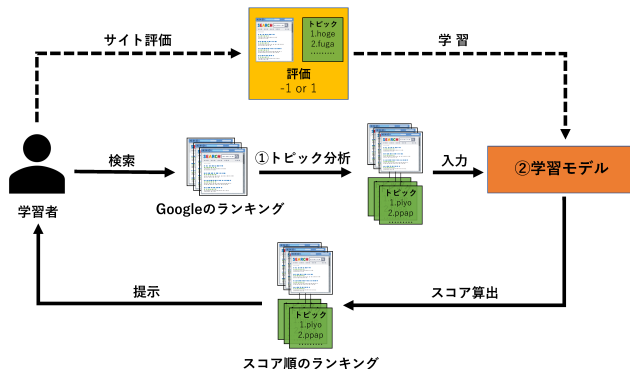


図1 提案システムの構成

まず、検索クエリに対する Google の検索結果 1 ページ目の Web サイトそれぞれにトピック分析を行い、サイト内のトピックを抽出する。次に、その抽出したトピックを学習モデルに入力することで、それぞれのサイトへの理解のし易さをスコアとして算出する。そのスコアの順でランキングを入れ替えて学習者に提示する。そして、学習者はサイトを閲覧して学習を行う。最後に、そのサイトに対して理解できたかどうかを 2 値で評価することで、そのサイト内のトピックに対する習熟度の更新を行う。

2.1 トピック分析

サイト内の文章にトピック分析を行うことでそのサイト内で扱われているトピックを抽出する。本研究では、Wikipedia のデータを基に PageRank の計算から文章内のトピックの抽出を行える、Wikifier[1] という Web サービスを利用する。

2.2 学習モデル

学習者のトピックごとの習熟度をモデル化し、サイトの理解し易さを予測するモデルとして、TrueLearn[2] を利用する。TrueLearn は TrueSkill というゲームのレーティングアルゴリズムを学習分野に応用したモデルであり、次のように置き換えることで、学習者のトピックごとの習熟度をモデル化する。

レーティング要素の置き換え

- プレイヤーのレート → トピックの習熟度
- チーム対戦 → 学習者の習熟度対サイト内のトピック
- 対戦の勝敗 → 学習者が理解できたかどうか

3 被験者実験

3.1 実験方法

被験者に提案手法を実装したシステムを利用して、Web 調べ学習を行ってもらった。実験する際に、被験者ごとに 5 回検索を行うこと、検索クエリは 1 単語にすること、検索する単語間は関連があるものにすることを条件とした。

3.2 評価指標

提案手法の有効性を評価するために、NDCG(Normalized Discounted Cumulative Gain), MRR(Mean Reciprocal Rank), MAP(Mean Average Precision) という 3 つの評価指標を用いる。これらの指標は値が 1 に近いほど、それぞれ、ランキング全体の理想的さ、正解ラベルのサイトが初めて現れる順位の高さのランキングごとの平均、正解ラベルのサイトの上位への集まり度合のランキングごと平均を表す。これらの 3 つの評価指標において、Google の検索ランキングと比較を行う。

3.3 実験結果

予備実験として、3 人の被験者に実験を行ってもらった。実験結果を評価指標ごとに平均すると結果は表 1 のようになった。

表 1 実験結果

評価指標	NDCG	MRR	MAP
Google	0.762	0.680	0.631
提案手法	0.816	0.852	0.682

3 つの評価指標全てにおいて、提案手法の方が上回る結果が得られた。特に MRR では顕著な差が現れ、ランキングの一位に学習者の理解し易いサイトがより現れていることがわかる。今後、実験回数と被験者数を増やした本実験を行うことで、提案手法と Google 検索における各評価指標の有意差の有無を検証する。

4 おわりに

学習者の検索履歴からトピックごとの習熟度のモデル化を行い、それに基づく理解のし易さからランキングの最適化を行った。そして、予備実験により、提案手法によって Google の検索結果よりも学習者により適したランキングを提示することができた。今後の課題としては、様々なサイトに対応することや検索結果の提示にかかる時間を短縮することなどが挙げられる。

参考文献

- [1] <http://wikifier.org/>,(参照 2020-12-27).
- [2] Bulathwela, S,Perez-Oriz, M,Yilmaz, E, and Shawe-Taylor, J. TrueLearn: A Family of Bayesian Algorithms to Match Lifelong Learners to Open Educational Resources. Proceedings of AAAI2020, Vol.34, pp.565-573, 2020. Artificial Intelligence.