

「なんとなく協調フィルタリング」手法の開発

近年、商品、学術文献、Webページなどが膨大に集積されるようになった。そのため、嗜好や要求に合うものを、利用者自身が、膨大なデータの中から見つけだすのが困難になってきている。そこで、利用者の代わりに、嗜好や要求に即したものを見つける推薦システムが開発された。

そのような推薦システムには、検索対象の特徴を利用する「内容に基づくフィルタリング」と、「口コミ」による推薦を自動化した「協調フィルタリング」の二つの手法がある。本研究では、後者の手法を改良した。

まず、協調フィルタリングの概略を述べる。この推薦システムには、いろいろな検索対象について、その好き嫌いを多くの人に尋ねた結果を集積したデータベースが事前に用意されているとする。推薦システムは、図1のような手順で利用者Aさんが好みそうなものを推薦する。①Aさんは幾つかの検索対象について、その好き嫌いを示すことで、嗜好を提示する。②システムは、Aさんと似た嗜好の人をデータベースから探す。③これらの人が好む検索対象を見つけて、Aさんに推薦する。口コミ情報と同様に、協調フィルタリングでは、自分と嗜好が似た人の推薦を参考にする。

この協調フィルタリングを実現するには「嗜好の度合いを測る」方法が必要になる。従来は、SD(Semantic Differential)法(図2(a))で嗜好を計測した。これは、個々の対象について、利用者に嗜好を5段階評価などで答えさせる方法である。しかし、SD法には、「好きさの度合いが利用者間で共有されている」などの不自然な仮定が必要になるという問題点がある。そこで、本研究では順位法(図2(b))を導入することで、この問題の改良を行った。この順位法とは、利用者に対象を好きなものから順に並べさせることで嗜好を測る方法である。順位法によって、より直観的に嗜好を提示できるようになるので、この協調フィルタリング手法を「なんとなく協調フィルタリング」と名付けた。

図3は、寿司についての嗜好を2つの推薦システムを用いて予測した結果である。横軸は利用者一人が嗜好を提示した対象の数、すなわち、何個の対象を評価したり、並べ換えたりしたかを表す。この個数が5個以上であれば、順位法の導入が予測精度を向上させていることがわかる。

今後は、さらに推薦の精度を向上させると共に、大規模なデータベースでも高速に実行できるような改良を行う予定である。

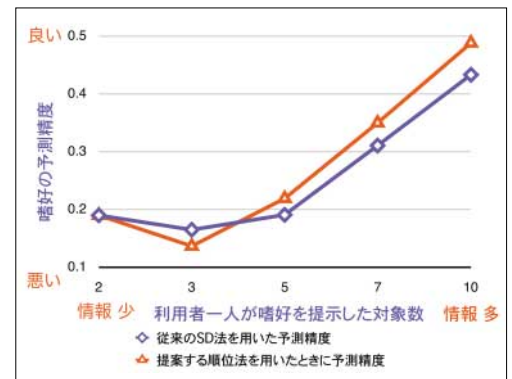
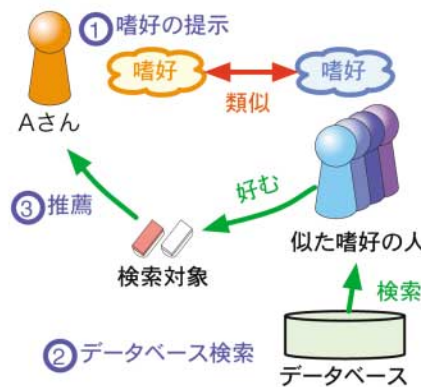


図1 (上) 協調フィルタリングの手順
図2 (右上) SD (Semantic Differential) 法と順位法
図3 (右下) SD法と順位法を用いた協調フィルタリングの予測精度



かみしまとしひろ
神島敏弘
mail@kamishima.net
脳神経情報研究部門

関連情報
● T. Kamishima, ACM KDD2003, 583-588 (2003).
● <http://www.kamishima.net/>
● 協調フィルタリングについて: <http://www.ai-gakkai.or.jp/jsai/whatsai/AItopics2.html>