

複合型フィルタリングによる図書推薦システムの開発

滝沢 伸也

これまでに図書館の貸出履歴を用いた図書推薦システムの研究が行われてきた。原田・増田の協調フィルタリングによる図書推薦を行うシステムの研究、貸出履歴に基づくアソシエーションルールにより図書推薦を行う辻らの研究がある。辻らの研究では、(a)辻らのアソシエーションルールによる図書推薦システム、(b)原田・増田の手法を再現した図書推薦システム、(c)Amazonによる図書推薦、の3つの手法から被験者に推薦を行った。それぞれの手法から推薦された図書に対する評価の比較を行った結果、(c)>(a)>(b)の順で推薦精度が高いということが確認された。

本研究では、辻らにより作成されたアソシエーションルールによる図書推薦システムを基に、図書の内容に基づく評価を加えた複合型フィルタリングによる図書推薦システムの作成を行い、その有効性の検証を行った。アソシエーションルールの作成には、T大学附属図書館の貸出履歴約6年分を用いる。図書の内容は、T大学附属図書館の蔵書データに含まれる書誌情報と、日外アソシエーツの「BOOK」データベースに含まれる図書の帯情報を用いる。複合型フィルタリングは、辻らのアソシエーションルールにより求められた『確信度』、『支持度』、蔵書データの『分類番号(NDC)』、『タイトル』、「BOOK」データベースの『帯情報』を属性としてSVM(サポートベクターマシン)で学習・識別を行い、その結果から推薦を行う。

複合型フィルタリングの有効性を検証するため、また、SVMのどの属性が有効であるかの検証を行うために(1)本研究で作成した複合型フィルタリングによる図書推薦に、(2)図書の内容(タイトルと分類番号)に基づく図書推薦、(3)タイトルの類似度による図書推薦、(4)辻らのアソシエーションルールによる図書推薦、(5)Amazonによる図書推薦の5つの推薦手法から被験者に対し推薦を行った。具体的には各手法が推薦図書として挙げる上位6件ずつを被験者に評価してもらった。被験者にはT大学に所属する大学院生を含む大学生で構成される27名を用いた。それぞれの手法から推薦された図書に対する評価の比較を行った結果、(5)>(3)>(4)>(1)・(2)という順に推薦精度が高かった。この結果から、複合型フィルタリングによる図書推薦システムの有効性は証明できなかった。その原因としてSVMの学習データが少なく、学習結果にアソシエーションルールの『確信度』と『支持度』、「BOOK」データベースの『帯情報』が適切に反映されなかったことが考えられる。(3)、(4)はそれぞれ(1)、(2)よりも良い推薦であるとの結果が得られた。また、(3)、(4)の推薦結果で一致している図書は1件も無かった。このことから、適切な学習を行うことが出来ればアソシエーションルールの結果と図書の内容に基づく評価からより良い推薦システムの作成が行えると考えられる。

(指導教員 辻慶太)