ファッションサイト向けの商品紹介文自動分類

Automatic Classification of the Product Description Sentence for Fashion Website

小川湧眞　黒澤義明　目良和也　竹澤寿幸

Yuuma Ogawa　Yoshiaki Kurosawa　Kazuya Mera　Toshiyuki Takezawa

広島市立大学大学院　情報科学研究科

Graduate School of Information Sciences, Hiroshima City University

1. はじめに

近年，市場の急速な成長を示しているオンラインショッピングにて，購買プロセスの省力化やユーザの購買意欲を煽る目的として情報推薦システムが用いられており，数多くの研究が行われてきた．

本研究では，オンラインショッピングのジャンルの中でも市場の大きいファッションに着目した情報推薦システムの一つとして，洋服の商品紹介文を洋服の構成要素に基づいて決定した属性に自動分類し，その属性を用いた推薦を行うことを考えている．過去にも属性分類を用いた洋服の推薦システムは提案されているが，提案されている属性の種類が乏しいという欠点があった．そこで，本研究では属性分類に商品紹介文を用いることでより高い表現能力を持った属性への自動分類を目的とする．

2. 提案手法

洋服の推薦に有用であると考えられる属性として，従来の研究で用いられている属性から 6 種類，商品紹介文の内容，アパレル関連の書籍から新たに 7 種類，計 13 種類の属性を提案した．本研究で用いる従来の属性と新たな属性を表 1 に示す．

表 1　提案した属性

新たな属性		従来の属性	
着心地	コーデ	季節	材質
目的	印象	シルエット	色
流行	効果	ディテール	柄
技術			

ファッションサイトの商品紹介文中の単語から，各文がどの属性に関する情報を有するのかを分類する．

13 属性それぞれについて SVM による機械学習を行って自動分類を行う．形態素解析を行い抽出した文中の名詞，動詞，形容詞，形容動詞の原形を素性，13 属性それぞれに属するのか属さないのかを正例，負例とする．

手法 1 では文中の全単語の頻度を素性として用いて学習を行う．手法 2 では，手法 1 の条件に加え， Web 上のファッション用語集[1]，同義語辞書から，各属性の分類に有用であると考えられる単語を収集，商品紹介文中に辞書内の単語が出現した場合に素性の重みを増加させ学習を行う．収集した各属性の辞書内の単語数を表 2 に示す．

表 2　各属性の辞書内の単語数

季節	シルエット	柄	効果
4	86	166	5
着心地	ディテール	技術	
5	607	42	
目的	材質	コーデ	
3	1,452	7	
流行	色	印象	
3	461	58	

3. 実験

学習データとして，ファッションサイトから収集した商品紹介文 1,200 文を用いて，13 属性それぞれについて属するか属さないかの二値分類を行った．分類器には TinySVM[2]を使用し, leave-one-out 法による検定を行った．

手法 1,2 における 13 属性の精度, 再現率の平均を表 3 に示す．本実験では，手法 2 における重み増加分は 10 倍に設定した．

表 3　実験結果

	精度(%)	再現率(%)
手法 1	42.9	44.1
手法 2	74.6	69.0

表 3 より, 手法 1 においては精度 42.9%, 再現率 44.1%, 手法 2 においては精度 74.6%，再現率 69.0%と値が上昇した．手法 2 において値が上昇した要因の例として属性「材質」における手法 1 の結果では,「オーガンジー」という属性「材質」に関する素性が含まれる文が誤分類されていたものの，手法 2 の結果では正しく分類された．この結果について「オーガンジー」という単語が 1,200 文中 1 文にしか出現しないため，素性の重みが低かったために手法 1 においては正しく分類されなかったためであると考えられる．しかし，手法 2 では「オーガンジー」の素性の重みを増加させたことで正しく分類されたのではないかと考えられる．

4. おわりに

本論文では, 商品紹介文を提案した 13 属性への自動分類を行うため，商品紹介文を形態素解析した結果を用いて機械学習を行った．実験の結果，手法 2 において 13 属性の精度，再現率の平均値として，精度 74.6%，再現率 69.0%という高い値を得ることができた．

---

[1] アパレルファッション wiki , http://apparelwiki.symphonic-net.com/

[2] TinySVM, http://chasen.org/~taku/software/TinySVM/.