

論文内容の要旨

情報工学コース

履 修 番 号	35315055	氏 名	JEON GOEUN
研 究 室 名	正田研究室		
研 究 題 名	レビューの類似度を利用した飲食店推薦		

論文内容の要旨

昔から食べ物の文化は、人類史の重要な部分を占めており、時代が変わると共に新しい食べ物が作られる現在、グルメサイトを通じて人々は簡単にお店と食べ物の情報を得ることができるようになった。例えば、好きな食べ物でお店を探すことや、近くの店で食べ物を選ぶ場合もある。また、グルメサイトで見られる他人のレビューで最も評判が良いお店を簡単に見つけることもできる。このように、人々はお店と食べ物の情報を事前に得て、お店に行ったことがなくてもそのお店について分かるようになった。だが、良い情報を短時間で得ることが可能であっても、自分自身がどんなものを食べたいのか決めてくれない。そこで、近くで簡単に見られるデータを用いて、食べたい食べ物があるお店を予想する推薦システムを考えた。

この研究では、食べログに登録されている口コミ、すなわち、他のユーザのレビューを利用してユーザ本人が食べたい食べ物があると思われるお店を推薦することを目的とする。食べログの様々なレビューを集めたデータと日本の食べ物用語辞書の食べ物の説明文のデータで類似性を計算するテキストマイニングを行う。類似性を求める手法としては、ベクトル空間モデルにおいて、単語同士の類似性を計算するコサイン類似度を利用する。また、自分自身を一つのユーザと考え、推薦リストを評価することで推薦が有効な推薦であるかを判断する。

本研究の推薦方法として、食べログの様々なレビューを集めたデータと日本の食べ物用語辞書の食べ物の説明文のデータを形態素解析エンジンである MeCab を用いて形態素解析し、それを用いて類似度を計算する。類似度の計算のため、抽出した単語を tf-idf 法で求め、その値をコサイン類似度の要素として類似度の計算をする。

その評価方法としては、それぞれの推薦結果を3段階で分け、最も正しいものは「2」、最も正しくないものは「0」、どちらでも当てはまらないものは「1」とし、各推薦方法の DCG を算出し、DCG の平均値が最も大きい推薦方法を今回の研究に適切な方法とする。

実験は3つの例、「そば」「すし」「ちゃんぽん」のレビューとキーワードで行い、レビューの類似度だけの推薦と比べた。

この実験の結果、レビューの類似度だけでは見られないお店の出現により意外性のある結果だと考えた。しかし、評価としては、「そば」は平均値が上がったため良い推薦に立っているとみられるが、「すし」と「ちゃんぽん」では、平均値が下がったり変わりがないため、良い推薦になっているとは判断しにくい。つまり、今回の推薦方法と例では適切な推薦になったのかは判断できない結果となった。