

自然言語処理とデータマイニングの研究

◆キーワード

自然言語処理 データマイニング 機械学習

◆産業界の相談に対応できる分野

自然言語処理 データマイニング 領域適応

工学部 情報工学科 講師

古宮 嘉那子

TEL 0294-38-5257

URL <http://komiya.cis.ibaraki.ac.jp/main.html>

e-mail kkomiya@mx.ibaraki.ac.jp

一言 アピール

本研究は、人間の言語理解を解明し、コンピュータによる言語の扱いやすさを向上させます。

研究概要

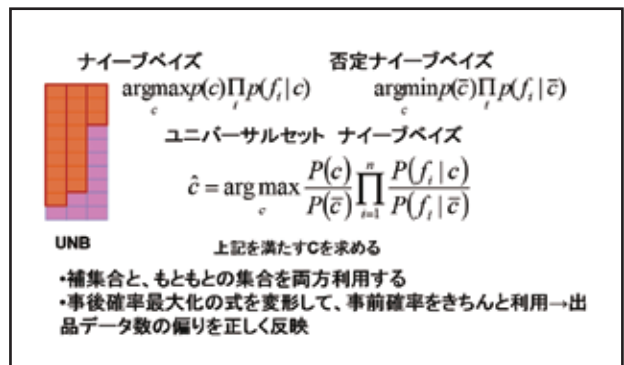
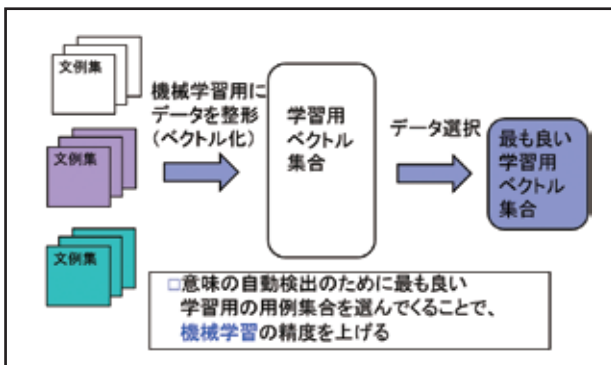
人間が言葉話すように、コンピュータにも言葉を理解させようという目的のもと、機械学習を用いて自然言語処理を行う研究が従来から数多くなされてきています。

言葉の使い方のルールの自動生成の研究もそのひとつです。たとえば、話者と聞き手の情報からの敬語の使い分け、人称代名詞、文末の機能語からの男言葉/女言葉の判定、また文脈からの語義曖昧性解消（言葉の意味の検出）を例に、言語選択に関するルールを自動的に作成しました。また、文中に現れる頻度から、言語選択に関係のありそうな原因表現を特定し、素性選択も含めてコーパスから直接行えるようにしました。

また、当研究室で力を入れている研究のひとつとして、語義曖昧性解消があります。テキストの入力に対して、そこで使われている言葉の意味を自動的に識別するものです。

通常の機械学習では、学習のデータと実行する対象のデータのジャンルが等しいのですが、当研究室では、テストデータと同領域の学習データがないという問題に対し、それらのデータの性質から使うべき領域適応手法を自動的に選択することにより、全体的な性能を上げることに成功しています。また、学習コーパスがないときに辞書を利用して言葉の意味を推定する研究も行っています。

また、ショッピングサイトにおける商品のカテゴリ分類や、テキストからのアニメ関係の用語の抽出、対訳対の抽出、日本語を理解しない外国人ユーザーに対する商品推薦などの研究を行ってきました。これらはオークションやネットショッピングにおいて、言葉の自動的な処理によって、ユーザーのニーズにこたえ、ユーザーを助けるシステムです。



何に 使える?

日本語を理解しない外国人ユーザーに対する商品推薦のようなことに利用・応用できます