

人工知能学会合同研究会 2015 に参加して

菊地 悠樹

Yuki KIKUCHI

電子情報学専攻修士課程 1年

1. はじめに

2015年11月12日(木)から13日(金)に、慶應義塾大学で開催された人工知能学会合同研究会2015に参加し、第106回知識ベースシステム研究会(SIG-KBS)において、「ソーシャルメディアにおけるユーザー行動にの基づいた料理レシピレコメンデーション」という題で20分の口頭発表を行った(図1)。



図1 発表の様子

2. 研究内容

2.1 研究背景

近年、Twitter や Facebook などのソーシャルメディアは、様々な情報を多くの人々が手軽に発信および共有できる媒体として進化しており、人々の日常生活から文化にまで大きな影響を与えはじめている。中でも最近、料理レシピ共有サイトがソーシャルメディアに登場し、普段の家庭料理のメニューを充実させるのに役立つとともに、人々の食生活を豊にすることにも貢献しつつある。このような背景から、料理レシピ共有サイトを対象としたデータマイ

ニングの研究が注目されはじめている。

料理レシピ共有サイトに料理レシピが投稿されたとき、それがどのユーザーに好まれるかを予測する手法の構築は、料理レシピレコメンデーションや人気料理レシピ予測等の観点から重要である。多くの料理レシピ共有サイトでは、他のユーザーが投稿した料理レシピに対して、ポジティブメッセージを与えることより、それに賛意を表明することができる。そこで、本研究では料理レシピ共有サイトにおいて、新たに投稿された料理レシピが将来、どのユーザーからポジティブメッセージを受け取るかを予測する問題を考える。

2.2 提案法

2.2.1 特徴量の抽出

料理レシピに対し、その文書データを形態素解析し、料理レシピを単語ベクトルにより BOW (bag-of-words) 表現する。次に、食材、調味料、調理法に関する用語の辞書を作成することにより、各単語が、食材であるか、調味料であるか、調理法であるか、またはどれでもないかを判別し、料理レシピを、食材単語ベクトルと調味料単語ベクトルと調理法単語ベクトルの組によって BOW 表現する。

2.2.2 レコメンデーション法

観測期間にユーザー達が投稿した料理レシピデータと与えたポジティブメッセージデータから、予測期間に投稿される料理レシピに将来、どのユーザーがポジティブメッセージを与えるかを予測する手法を提案する。

ユーザーが観測期間に投稿した料理レシピおよびポジティブメッセージを与えた料理レシピと、予測期間に投稿された料理レシピとの類似度をそれぞれ、コサイン類似度で測定し、予測する料理レシピに対するユーザーのスコアをすべてのコサイン類似度の最大値で定義する。そのスコアの値に従って、料理レシピに将来ポジティブメッセージを与えるユーザーを予測する。

2.3 評価実験

料理レシピ共有サイトの実データを用いて実験を行った。提案法を cos+_{max} とし、その性能を評価するために、次の5つの手法と性能を比較した。第1の手法は、 cos+_{max} 法ではユーザのスコアを類似度の最大値で定義したものを、類似度の平均値に変更したものである。この手法を $\text{cos+}_{\text{mean}}$ と表わす。第2の手法は、 cos+_{max} 法では料理レシピ間の類似度を測定する際、料理レシピを食材単語ベクトルと調味料単語ベクトルと調理法単語ベクトルの組で表現したが、それを単純に単語ベクトルで表現するように変更したものである。この手法を cos_{max} と表わす。第3の手法は、 $\text{cos+}_{\text{mean}}$ 法において、第2の手法と同様、料理レシピ間の類似度を測定する際、料理レシピを単純に単語ベクトルで表現するように変更したものである。この手法を cos_{mean} と表わす。第4と第5の手法は、確率的文書生成モデルである HDP-LDA を用いるものである。第4の手法では、各料理レシピを食材単語ベクトルと調味料単語ベクトルと調理法単語ベクトルの組で表現し、第5の手法では、各料理レシピを単純に単語ベクトルで表現する。どちらの手法も、観測期間と予測期間内のすべての料理レシピデータを HDP-LDA に学習させて、各料理レシピのトピックベクトルを推定する。そして、ユーザが観測期間に投稿した料理レシピおよびポジティブメッセージを与えた料理レシピから、予測期間に投稿された料理レシピへの距離をそれぞれ、カルバックライブラー情報量で測定し、ユーザのスコアをすべての距離の最小値で定義する。スコアの値に従って、料理レシピに将来ポジティブメッセージを与えるユーザを予測する。第4の手法を HDP-LDA+ と表わし、第5の手法を HDP-LDA で表わす。

評価結果を図2と図3に示す。これらの結果より、料理レシピの特徴量としては、食材単語ベクトルと調味料単語ベクトルと調理法単語ベクトルの組を用いることが有効であること、また、新たに投稿

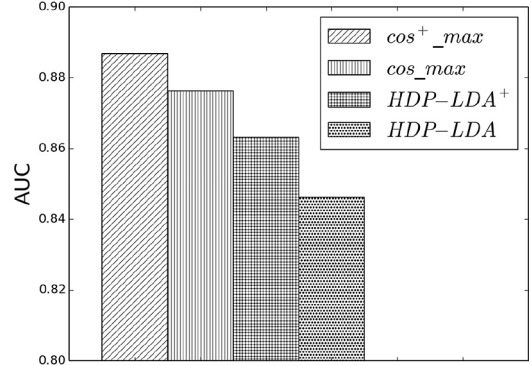


図2 提案法と他手法の予測性能比較

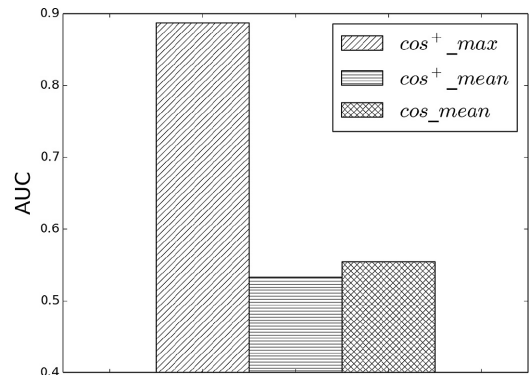


図3 提案法と他手法の予測性能比較

された料理レシピに対するユーザのスコアとしては、料理レシピ間の類似度の平均値ではなく最大値を用いることが非常に有効であること、さらに、料理レシピの間の類似度としては、複雑な確率モデルではなく単純なコサイン類似度を用いることが有効であることがわかった。

3. おわりに

学会では様々な研究発表を聞くことができ、また本研究に多くの意見をいただき貴重な経験となりました。今後の研究に生かしたいと思います。

最後に、今回の発表にあたり、ご指導を頂いた木村昌弘教授、熊野雅仁実験講師にこの場を借りてお礼を申し上げます。