

## 平成30年度 情報工学コース卒業研究報告要旨

|  |                          |       |
|--|--------------------------|-------|
| 武田(浩)研究室   | 氏名                       | 山田 康輔 |
| 卒業研究題目   | SNSにおける行動類似性を用いたユーザの性格推定 |       |
| <p>本研究では、SNS ユーザを対象とした性格推定に取り組む。従来の性格推定に関する研究では、ユーザの投稿したテキストやフォロー数のような統計情報を利用したものを中心に行われてきた。しかし、SNS ユーザは投稿だけではなく、「いいね」や「シェア」のような投稿に対する行動も多く行っており、これらの情報はユーザの嗜好や共感を表し、性格推定において重要な手がかりになると考えられる。たとえば、性格特性が外向型であるユーザと内向型であるユーザによく「いいね」をされた投稿に注目すると、それぞれの投稿の内容には、「いいね」をしたユーザの性格特性が反映されることが考えられる。実際に収集したデータを分析したところ、対人関係に積極的な外向型のユーザは実際に行われている交流イベントに関する投稿によく「いいね」をし、対人関係に控えめな内向型のユーザはイラスト付きの投稿によく「いいね」をすることが確認された。そこで、同一の性格特性を持つユーザはよく「いいね」をする投稿が類似すると仮定し、同一の性格特性を持つ他ユーザとの「いいね」や「シェア」の類似性を利用した性格推定を行う。</p> <p>本研究では、Twitter を使用しているユーザを対象に性格推定を行い、推定する性格特性には Myers-Briggs Type Indicator (MBTI) という性格検査により診断された性格タイプを利用する。MBTI は性格特性に関する 4 つの指標においてそれぞれ 2 つの型に分ける 16 タイプに分類する検査であり、4 つの指標に対する二値分類問題として性格推定を行う。分類器には教師あり学習の代表的なモデルの一つであるサポートベクターマシン (SVM) を利用する。SVM の素性には、従来用いられてきた紹介文とツイートからなるテキスト情報と統計情報による素性 (素性集合 1) に加え、同一の性格特性を持つ他ユーザとの「いいね」や「シェア」の類似性を利用した情報による素性 (素性集合 2) を用いる。それぞれの素性集合によるモデルを構築し、それら 2 つのモデルを結合させた性格推定システムを最終的なモデルとして提案する。</p> <p>実験には 60 個の質問に基づき MBTI の検査を行う Web サイトで診断された結果をツイートした 72,847 人の Twitter ユーザの性格タイプを用いて、提案するモデルの評価を行った。2 つのモデルを結合させたシステムと素性集合 1 のみからなるモデルとの精度の比較を行ったところ、利用するツイート数が多い場合では素性集合 2 を用いることによる効果はあまり確認できなかったが、利用するツイート数が少ない場合では推定精度の向上が確認できた。このことから提案する手法では、投稿したツイート数は少ないが「いいね」や「シェア」を頻繁に行うユーザへの性格推定に特に有用であるといえる。</p> |                          |       |