

顔画像を用いた性別予測と SNS 上での行動予測

プロジェクトマネジメントコース 矢吹研究室 1442045 川辺明俊

1. 序論

ネットリテラシーとは、情報ネットワークを正しく利用することができる能力のことである。リテラシーとは、もともとは識字能力のことで、文字や言語に対する能力の意味である。それに「ネット」と付け加えることで、インターネットを使いこなす基本的な能力を指す言葉として「ネットリテラシー」が定着した [1]。このネットリテラシーが不足していると、インターネットを使用する際に、不正サイトでクレジットカード番号を盗まれたり、コンピューターウイルスの感染により、個人や会社などの情報を流出されたり、ネット上にある嘘の情報に騙されてしまう。実際に、熊本地震でライオンが動物園から脱走したと、デマ情報を流し Twitter 上で拡散した男性は、偽計業務妨害容疑で逮捕された。

Twitter とは、インターネット上で「ツイート」と呼ばれる 140 文字以内のメッセージや、画像、動画、URL を投稿できる情報サービスである。日本では、現在 (2017 年 10 月) 利用者が 4,500 万人にもなるソーシャル・ネットワーキング・サービス (SNS) と見られている。

Twitter では日々、大量の情報がツイートされる。もちろんデマ情報などの、ネットリテラシーを問われるような、情報も錯綜している。そこで私は、Twitter のデマ情報に騙され、ネットリテラシーが不足している人の情報を解析し、分析できるのではないかと考えた。

2. 目的

Twitter のデマ情報を信じ、情報を拡散してしまうのに男女間で差は生まれるのかを調べる。

3. 手法

研究方法は以下のとおりである。

1. 機械学習で男女の性別を、判定できるように wikipedia のプロフィール画像をもとに、Neural Network Console を使用し、男女の顔画像を判

別するための学習済みモデルを作成する。

2. デマ情報のツイートをリツイートした人のプロフィール画像を集める。
3. 集めたプロフィール画像を、始めに作成した学習済みモデルを使用し、機械学習で性別を判定させる。
4. 男女の数を集計し、どのぐらい差が生じるか調べ、考察する。

4. 結果

本研究の結果として、Neural Network Console を使用し、男性と女性の判別をした場合、80 %程度の精度しか出なかった。デマ情報を拡散する Twitter のユーザーと、正確な情報を拡散するユーザーには関わらず、男性と女性では差あるのかは、分がなかった。

5. 考察

Twitter ユーザーのプロフィール画像は、自分の顔画像を使用していることはとても少なく、データが不足してしまうことが分かった。そして、写真から男性と女性を学習させた学習済みモデルで、画像認識を使用した場合、機械学習の性能が低いため、性別を判別するのは、難しいと考えた。

6. 結論

本研究ではネットリテラシーの不足している人には、どのような特徴があるのかを調べた。だが、Twitter では、プロフィール画像で自分の写真を使用しているのはわずかであり、機械学習を使用した画像認識では、性別を判別するのは難しいことだと分かった。

参考文献

- [1] ネットリテラシー -インターネット用語辞典- | OCN. <http://www.ocn.ne.jp/support/words/online/83l83b83g838A83e838983V815B.html>.