

オンライン小説の検索に有効なタグの推薦に関する研究

山崎睦月

近年、オンライン小説に特化した CGM サイトである小説投稿サイトは、投稿が容易で、様々な小説がいつでも無料で利用できるため、多くのユーザに利用されている。しかし、投稿小説が増加することによって、投稿者と読者、それぞれのサイト利用に問題が発生している。投稿者は、投稿した小説がほかの小説に埋もれてしまい、読者の目に触れる機会が少なくなっている。読者は多くの小説の中から自身の趣向にあった小説を探すことが難しくなっている。

小説投稿サイトでは、利用者（主に投稿者）がタグを小説に設定することができる。タグを設定することで、タイトルやあらすじだけでは表すことのできない情報を作品に付与することが可能となっている。付与された情報によって、小説は検索されやすくなり、ニーズのあった読者の目に触れる機会が多くなると考えられる。しかし、タグは投稿者に依存するため、検索の絞り込みが難しいタグを設定したり、意味の重複したタグを設定したりしてしまっている。このように、タグが有効活用されていないケースも多く見受けられる。

本研究では検索を支援するという視点から、絞り込みに効果的なタグを小説投稿者に推薦する手法について提案する。ここで、有効なタグとは、小説の内容を合致しているかつタグ間で意味の重複がない、小説の数に対して1つのタグが小説に設定された数（タグの出現数）が十分に小さいタグである。まず、タグ全体からタグの出現数など参考にして、検索に有効なタグの範囲を求め、タグの抽出を行う。それから、小説本文を Doc2Vec を用いてベクトル化を行う。次にタグのベクトル化も行う。タグは設定された小説の特徴を表すものであるため、タグは設定された小説と同様な性質を持っていると考えられる。したがって、タグは設定された小説のベクトルを合成したベクトルを持っていると考えられる。推薦対象の投稿者が入力した小説本文もベクトル化し、タグのベクトルと類似度比較を行い、タグの推薦を実行する。この時、推薦するタグの中で意味が重複することを防ぐために、タグ間でも類似度の比較を行う。

(指導教員 佐藤 哲司)