Twitterデータを用いた政治的論点の抽出と政治家分類

野村　真由美

要旨

日本の国政選挙において，特に若い世代を中心にその投票率の低さが問題となっている．そのような中で，近年野党の新党結成や合併等が多くみられ，その都度人の出入りが激しい政党が現れていくこととなった．それに伴い，政党の乱立において論点や立場が判然としない団体が増えていることが予想される．さらに，政治広報や選挙公報に活動や履歴の全てが記載されるわけではなく，政治家の状態や思想が一目でわかる資料はなかなかに見つけにくい実情がある．また，政治的論点は時が経つにつれて変化し，同じ問題であっても表現に使われる言葉もまた変化をしていく．さらに，そもそも政党はただ1つの公約だけで選挙を戦うということはしない．そのため，すべての言葉は複数の論点が入り混ざることによって，その政治家ないし政党の立場を示すものとなる．現在，論点を1点与ることによって政治家やツイートを分類する研究が先行して行われている．しかし，時が経ち変化していく複数の論点を基に，議員個人の政策がつくられ選挙を戦うことになることを考慮する必要があると考える．このとき，現状存在する特定の政治的論点1点からのみの分析だけでは不十分であり，各論点において複合的に分類し，かつ論点の発生を検知できるシステムが必要となる．そこで，政党所属議員や近しい人間が発信するTwitterをデータに用いて，政治公約等から見えない議論や選挙モードでない通常の政治家の発信に着目する．この発信から，政治家の実情を伝え，政治的論点や政治的諸問題，そしてその立場が分かりやすく判別できるような投票活動につなげるシステム作成を行うことを目指す．その一環として，このデータより複数の論点の抽出と，その論点による政治家の分類を行う．本研究では，政治関係者のツイートより、単純な発信数の多い単語と，係り受け解析によって賛否両論が起きる単語を，それぞれ抽出した．これを基に，それぞれの抽出単語より人と単語のベクトル表現を使用し，発した人間をクラスタリングで分類する．また，抽出された言葉同士の関連性にも着目し，言葉の立場の分類を行う．