

地理情報を用いた多面的・多角的な思考支援システムの構築

川崎 誠也

近年の急速な情報化の発展によりデータの集合から意味を見出し、正しく解釈する能力（データリテラシー）が求められるようになった。データ収集から分析までのプロセスを実践的に学ぶデータサイエンス教育の需要が高まっている。また、防災、SDGs（持続可能な開発目標）の観点から地理教育が求められている。

そこで、地理情報を用いてデータの様々な分析方法を学び多面的・多角的な考察、表現を補助するシステムを開発した。本研究では、オープンデータを利用して作成したグラフや解析例とそこから読み解けることの解説をつけ、位置情報と複数のデータから地域のつながりや、事象間の関連を見つける力を養うことを目指した。

データから知見の獲得までのプロセスを KDD (Knowledge Discovery in Databases) プロセスに基づき、①選択、②可視化、③解析、④結果の可視化、⑤解釈・評価の5段階に分け、データ分析の一連の流れを学習できるシステムを開発した。都道府県ごとの社会生活に関するオープンデータを利用し、Web ブラウザ上でインタラクティブに学習をすすめていくことができる。主成分分析やクラスタリングといった手法を用いて解析した結果を地図上で可視化した結果等から学習者は解析手法を学びデータを読み解く力を養う。

学習者は、1巡目は項目数を2つで行う単回帰分析、2巡目は3つ以上の項目で行う重回帰分析という流れで進める。項目数が増えると特徴を可視化して考えづらくなるため、クラスタリングやその前処理である標準化等が必要となる。各段階で適切な手法や評価方法と注意点を示す。一連のデータ分析を学習した後、解説のない状態で新たなデータと、その変換されたグラフを閲覧して被験者自身が何を読み取れるか考えることで、実際のデータ分析でどのようなアプローチができるか検討する機会を提供するシステムを実現した。

システムの有意性を評価するため、大学生7人を対象に実験を行った。システム利用前後にアンケート及びインタビューを実施した。インタビュー調査で、インタラクティブに試行できることが理解の促進につながった、人によってデータの解釈が違うことに興味を持ったという意見があった。このことから複数の視点からデータを読み解くという目的は達成され、KDD プロセスにそって繰り返しシステムを利用することで、複数の考え方や批判的な見方につながることがわかった。一方、データ分析の一連の流れを学習したものの、別のデータに対して学習者自身が新たな分析を始めるには不十分であったと考える。親しみをもちやすいデータを用いたことや、計算式を省き全体の流れを重視したこと、グラフや解析結果の読み取りに解説を加えたことによりデータ分析に対して興味をもった被験者が多かった。

今後、学習者がデータを選び可視化した結果をシステム上に蓄積していき、各々が読み取ったことを他者と共有することでさらにデータの多面的な見方を増やすことが期待できる。

（指導教員 時井真紀）