

特集にあたって

生田目 崇 (中央大学)

本号は、経営科学系研究部会連合協議会主催「令和3年度データ解析コンペティション」(以下、「本コンペティション」)の研究成果をもとに査読付き論文を募集し、採録された論文により特集を組まさせていただきました。6編の論文が投稿され、いずれも2名の査読者による査読を経て、4編が採録されました。毎年のことになりますが、短期間で査読をお願いするようになりました査読者の皆様には感謝の念に堪えません。

本コンペティションでは、コネヒト株式会社様より同社が運営する妊娠・出産・子育てをしている女性をターゲットとしたコミュニティサイト「ママリ」のデータを提供いただきました。提供いただいたデータは、

- ユーザ (母親) 情報 (ユーザ ID, 居住地)
- 子供情報 (ID, 誕生月, 性別)
- 質問 (投稿者 ID, カテゴリ, 質問テキスト)
- 回答 (回答者 ID, 質問 ID, 親回答 ID (回答に対するコメントの場合), 回答テキスト, 「いいね」の有無)
- 検索 (検索者 ID, 検索ワード)

であり、データ期間は2019年1月～2021年7月で、提供データは大量のテキストが含まれていることが特徴です。データサイズは概算で、20 Gbyte、ユーザ数300万人、質問と回答合わせて4,000万件、検索回数はおおよそ1億2,000万件という規模のデータでした。貴重なデータを提供いただいたコネヒト株式会社様に御礼申し上げます。

本コンペティションの成果発表会では、今回の対象がターゲットの絞られたコミュニティサイトであり、データの多くがテキストということもあったので、テキストのコンテンツの分析が多くを占めるかと思いましたが、意外にも定量的な分析も少なくなく、研究の幅は大変広がったのが印象的でした。テキストは一つ

の情報として、ユーザのプロファイリングなど生活者の状況に切り込んだ研究や、子供の月齢に着目した研究、質問を軸にした回答のつながりなど、大変興味深い成果も数多くありました。

子育てを経験された方はおわかりと思いますが、妊娠から出産、小さい頃の子育ては、ライフステージが月ごとに変わっていき、特に初めての子供のときはすべてが初めてのことで、疑問や悩みがつきません。出産が近づくもしくは子供が成長するにつれ次々とそれまでの日常生活では見慣れないような出来事に頭を抱える日々を過ごすことも多いでしょう。都市化、核家族化が進み、また子供の数も減っている現在においては、悩みを相談したり情報を共有する場が減っていると思われませんが、こうしたコミュニティを介して同じ心境のママとのつながりや、先輩ママや母親世代のユーザから解決策をいただけるなど、コミュニティサイトは新米ママにとっても大変心強いものと思われま。こうしたサイトのデータを分析できたことは、もちろん自然言語処理の対象として望まれるというだけでなく、生活者の問題解決やコミュニティの理解などさまざまな面からの研究のきっかけになったものと思います。今回掲載された論文をぜひ一読いただければ幸いです。

データ分析環境については株式会社 NTT データ処理システム様にツールの貸与をいただきました。毎回快く貸与いただくことに感謝いたします。

今年度(令和4年度)のコンペティションでは、楽天グループ株式会社様より、新型コロナウイルス感染症前後の楽天市場の購買履歴データを提供いただいています。今年度こそは対面での成果報告会開催ができればと思っております。