

データ利活用促進に向けたデータ分析に必要なデータを効率的に収集する技法の研究(クラス 1)

アブストラクト

1. 背景

現在、あらゆる産業でデジタルトランスフォーメーション（以下 DX）への期待が高まっている。DX の早期実現に向け、AI/IoT 等のキーテクノロジーを有効活用したデータ利活用を推進していくことが求められており、AI の市場規模は年々拡大傾向にあることから AI への期待の高さが伺える。

DX を成功に導くデータ利活用の一環として、AI を活用したデータ分析が挙げられる。しかしながら、データ利活用の導入実績では、日本は米国・ドイツから遅れている。その起点であるデータ収集の段階から遅れをとっていることを踏まえると、データ利活用において、良好なデータ分析結果が得られるデータ収集技法を確立することはデータ利活用の促進に寄与する役割は大きいと言える。

2. 研究目的

本分科会では、AI 分析において良好なデータ分析結果の得られるデータ収集技法を提案することを研究目的とする。具体的には、分析事例別に最適な分析アルゴリズムの選定、ならびに収集データとデータ分析結果との関係性に着目した良好なデータ分析結果が得られるデータ収集技法の提案を行う。

3. 研究のアプローチ

本分科会では、効率的なデータ収集技法を AI・機械学習の分析アルゴリズムを対象として「実例調査」と「自己検証」でデータ収集の知見を調査し、結果をデータ収集ガイドラインにまとめた。「実例調査」とは、AI 分析のビジネス事例を調査しデータ収集技法を整理したものを指す。一方、「自己検証」とは分析アルゴリズムを用いてデータ分析を試行し、データ収集技法を導き出したものを指す。

4. データ収集ガイドライン

データ収集ガイドラインは、本分科会で実例調査・自己検証を実施した結果、得られた知見についてまとめたドキュメントである。データ収集ガイドラインは図 1 に示すように 3 つのシートから構成されており、それらを活用することで適用先に合った機械学習のアルゴリズムとデータ収集技法の提案ができる。本ガイドラインの利用の流れは図 1 の通りである。

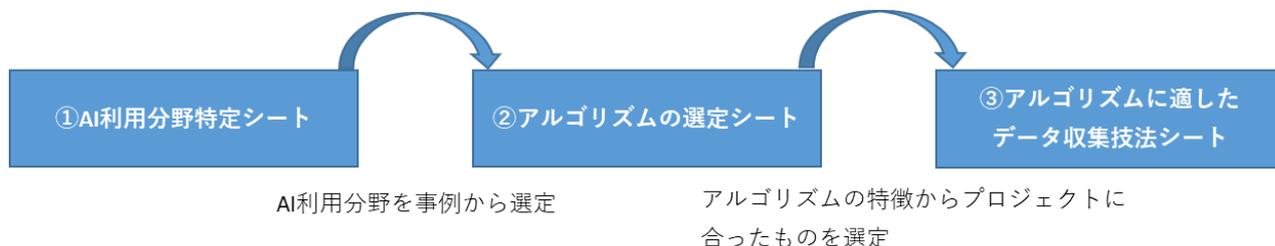


図 1：データ収集ガイドラインの構成と利用の流れ

5. 総括

本分科会では研究目的を達成するために、機械学習の分析アルゴリズム毎に効率的なデータ収集技法や機械学習を利用する際のデータ例をデータ収集ガイドラインにまとめた。また、本ガイドラインは適用先にマッチするビジネス事例から分析アルゴリズム・効率的なデータ収集技法が選択できる点で有用であり、本ガイドラインがデータ利活用の促進に寄与することを期待したい。