

インフラ維持管理データサイエンスの体系化と試行支援

北海道大学 正会員 ○湧田 雄基
 東京大学 正会員 石川 雄章
 ファイブソリューションズ株式会社 磯部 正幸

1. 緒言

我々は、土木インフラの維持管理業務を対象に、データサイエンスの活用に関する研究を進めてきた¹⁾。データサイエンスは、日本の土木分野において大きな注目を浴びているが、実践には試行錯誤の時間を要する場合が多く、また分野特有の課題や、その解決方法が存在する。そこで、本研究では、これまでのインフラ維持管理分野におけるデータサイエンス事例およびデータ準備等のノウハウを体系化し、これを用いてデータサイエンスの効率を向上し、試行サイクルを早くするための「インフラ維持管理のためのデータサイエンスクラウドサービス」の開発を実施してきた。本報では、クラウド設計に反映したインフラ維持管理分野におけるデータサイエンス試行の体系化と支援クラウドサービスのプロトタイプについて報告する。

2. 手法

インフラ維持管理業務では、データや根拠に基づいた安全のための判断ということが重要な要件であり、AIやデータサイエンスとの親和性は高いと考えられる。ただし、著者らは、「AIが何でも解決してくれる」といった取り組み姿勢では、インフラ分野での実用的な活用には至らないと考えている。本研究でのデータ活用の狙いは、AIやデータ分析処理と人とが連携しながら、業務に貢献可能な導入を進める事である。AI等が得意とする作業が何かを十分に整理した上で、これを最大限活用し、維持管理業務に従事する人の「データやAIに基づいた論拠有る意思決定」の支援を行い、業務の効率化を目指す。このためにも、インフラ維持管理従事者に対し、高度なデータ活用としての「データサイエンス」を体験し、データやAIに対する理解を深めるとともに、業務に活用するためのデータサイエンス試行支援を行う仕組みの構築を目指している。

データサイエンスの流れに関して、本研究では図-2の構成を提案している。手順構成（プロセス）は、次の6つとして考えている：現場・業務理解、データ理解、データ準備、分析、評価、現場への展開。この流れは、一般的なデータマイニングにおけるプロセスを規定するCRISP-DM（Cross-Industry Standard Process for Data Mining）を元に、インフラ維持管理のデータサイエンスに適すようカスタマイズしたものである。

データサイエンスの初頭を実施する「現場・業務理解」は、インフラ維持管理における特徴的なプロセスと言える。この部分は一般には、Business Understanding（ビジネス理解）とされるプロセスであるが、インフラ維持管理分野におけるデータサイエンスのモチベーションは、経営目標とは限らず、「維持管理現場の業務/労務上の効率向上」や、あるいは、「現場の維持管理の質向上につながる知見を得る事」がデータサイエンスのゴールになり得る。

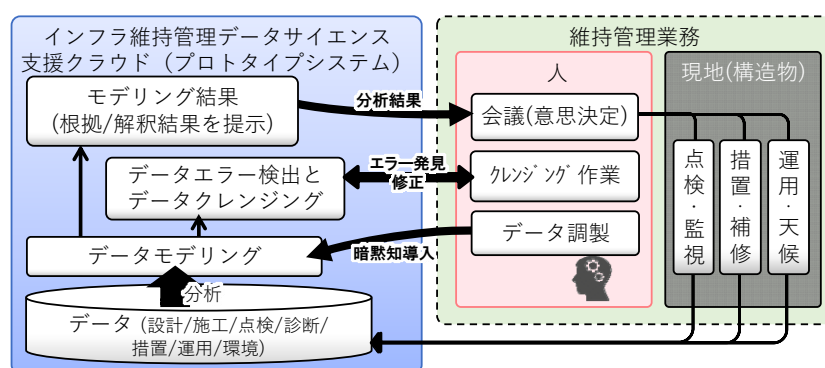


図-1 インフラ維持管理データサイエンスにおけるクラウドの利用

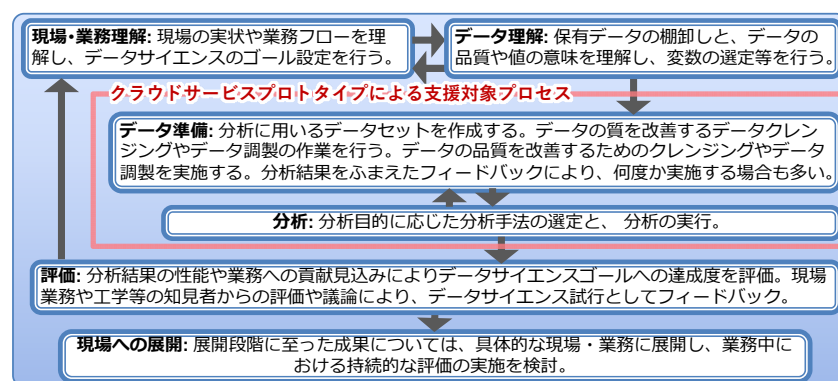


図-2 インフラ維持管理におけるデータサイエンスの手順

キーワード データサイエンス, 機械学習, データマイニング, インフラ維持管理
 連絡先 〒060-0814 札幌市北区北14条西9丁目 北海道大学 数理・データサイエンス教育研究センター
 E-mail: wakuda@lmd.ist.hokudai.ac.jp, TEL: 011-706-4401

我々は、これまでに、前述のデータサイエンスのプロセスおよび、データクレンジングやデータ調製の手法に関する過去事例の整理に基づき、データサイエンスを実践するためのクラウドサービスとして、インフラ維持管理データサイエンス支援機能（以降、システムという）のプロトタイプ開発を行った²⁾。

提案するシステムは、これまで著者らのチームが実施してきたデータサイエンス研究で蓄積してきたノウハウや要望をふまえ、そこで出た要望に対応し、データサイエンスを効率よく実施するために開発を行った。従って、複数人での同時に分散的なデータ準備作業を行えるように、また、データや分析結果を共有できるように、クラウド上のサービスとして設計・実装した。また、何でもできる汎用性を提供するよりも、インフラ維持管理分野のデータにターゲットを絞り、データサイエンス試行の際に踏むべき手順を適切にサジェストし、ユーザが迷わず操作が出来るツールを目指している。

提案するシステムの重要な特徴は、ドメイン知識に基づいた変数の知識や、データエラーの辞書、およびデータ調製の手続きを辞書として保持している事であり、これらの知識データベースは過去の我々の研究事例から予め作成してあるとともに、システムを使用する度に蓄積される。これにより、ユーザは、登録したデータや目的に応じ、有効な分析をいち早く試すことができる点が大きな特長である。

データクレンジングやデータ調製については、多くの場合には、手続きが煩雑化するとともに、実施した処理の記録や再処理、あるいは元に戻す処理が必要となる。提案クラウドでは、これらのデータクレンジングやデータ調製、分析等を実施した際の処理の手続きを「レシピ」と呼び、知識データベースとして管理する。この知識データベースによる試行のサポートが、データサイエンスを円滑に実施するために重要な技術である。これによって、あたかも専門的なデータサイエンティストが行うような、データ調製や分析手法の選定と実行を試行可能となり、またユーザ同士で共有できるようになる。このようなデータサイエンスにおける操作の流れであるレシピは、形式が類似する他のデータセットに対しても適用が可能であり、典型的なデータ処理・分析手続きの流れを他の構造物に適用したり、他者と共有することも可能となる。

No.	カラム名	種別	レコード数	非空白件数	有効率	ユニーク件数	基キタイプ件数	要約統計量(数値のみ対象)	頻出値 #1	頻出値 #2
1	年度	数値	12,076	12,076	100.0%	5	0	12,076	0	2016
2	年度	数値	12,076	11,913	98.7%	87	163	11,913	0	89
3	年度	数値	12,076	12,076	100.0%	746	0	12,076	0	89
4	年度	数値	12,076	11,918	98.7%	10	158	11,918	0	10
5	年度	数値	12,076	11,768	97.4%	7	266	11,768	0	10
6	年度	数値	12,076	10,838	89.7%	883	1,238	10,838	0	10
7	年度	数値	12,076	12,044	99.7%	383	32	12,044	0	10
8	年度	数値	12,076	12,042	99.7%	380	34	12,042	0	10
9	年度	数値	12,076	11,853	98.2%	983	223	11,853	0	283
10	年度	数値	12,076	12,076	100.0%	504	0	12,076	0	10
11	年度	数値	12,076	11,785	97.6%	4	291	11,785	0	10
12	年度	数値	12,076	11,589	96.0%	23	487	11,589	0	10
13	年度	数値	12,076	12,076	100.0%	7	0	12,076	0	10
14	年度	数値	12,076	12,044	99.7%	452	12	12,044	0	10

図-3 データサイエンスクラウドサービス画面（開発中画面）

システムのプロトタイプは、限定的に研究関係者で利用し、評価を実施中である。ユーザにより、点検や構造物諸元、措置記録等のデータテーブルをアップロードし、データの内容確認（図-3 は、開発中画面）や、データクレンジング、データ調製、分析が可能であり、データサイエンスのプロセスにおいて、試行運用を行っている。試行の結果をふまえ、応答性や GUI 等の工夫を適用し、利用感向上を行って居る。今後、分析手法の拡充や動作の安定性を高め、希望利用者への公開を行う計画である。

3. 結言

本報では、著者らが実施してきたデータサイエンス事例を整理することで、インフラ維持管理におけるデータサイエンスの手順の体系化を行い、これに基づいたデータサイエンスの試行支援を行うためのクラウドサービスのプロトタイプについて報告を行った。今後は、データサイエンス事例に基づくレコメンデーション機能など、さらなる利便性の向上と、データサイエンスサイクルのスピードアップに貢献するよう、事例の整理および機能の拡充設計・実装を進める。

謝辞

本研究開発の一部は、総務省 SCOPE #162103120 の委託および、JSPS 科研費 JP00377847 の助成を受けたものです。

参考文献

- 1) 東京大学 大学院情報学環 社会連携講座「情報技術によるインフラ高度化」, <http://www.advanced-infra.org/english.html>, accessed on January 29, 2018.
- 2) Wakuda, Y., Isobe, M., and Ishikawa, Y., Data Science Support System for Infrastructure Maintenance on Cloud, 17th International Conference on Computing in Civil and Building Engineering (ICCCBE 2018), 2018.