

AI ビッグデータ応用における仮説解釈説明の構造探求と関連情報通信技術の体系化

代表研究者 石川 博 東京都立大学 システムデザイン研究科 特別先導教授
共同研究者 遠藤 雅樹 職業能力開発総合大学校基盤ものづくり系情報通信准教授

1 研究調査の概要

本調査研究では、仮説（モデル）の解釈の構造とその構成要素となる説明の種類を明らかにして、その構造に基づいて、現存する個別の解釈説明の情報通信技術の体系化を図る。仮説の解釈説明に適した仮説の記述手段を探求し、それを提案する。体系化された解釈説明の要素の中で、現状では欠落または不十分な要素を明確にして、それに対する有望なアプローチを探求し、それを提案する。以上の提案検証のために多様なユースケースを収集する。

2 研究調査の詳細

2-1 研究調査の目的

現在活発に推進されている AI（人工知能）や機械学習の実用化、ビッグデータ応用システムの受容において仮説（モデル）の解釈説明が必要とされている。

【リサーチクエスチョン】

将来の Society 5.0 [参考文献 1] 実現上でカギとなる AI やビッグデータが広く受け入れられ利用されるようにするためには、その応用システムの利用者（分野専門家や一般ユーザ）に対して、マイクロな説明（仮説生成の説明とモデルの決定に対する理由の説明）及びマクロな説明（仮説生成を含む応用システムの全プロセスの説明）の両方を行う必要があるのではないか、またそのためにどんな基盤技術を開発すべきかというリサーチクエスチョンを解決する必要がある。さらに利用者が個人情報に関する説明を納得できると個人情報の活用を促進する情報銀行 [参考文献 2] や個人情報保護に関する EU の GDPR 法 [参考文献 3] に基づくデータポータビリティの進展も加速される。

【研究目的】

そのために本申請では研究目的を以下の項目に分けて研究を実施する。

- ・ 仮説の解釈の構造とその構成要素となる説明の種類を明らかにする。
- ・ 仮説の解釈説明の構造に基づいて、現存する個別の情報通信技術の体系化を図る。
- ・ 仮説の解釈説明に適した仮説の記述手段を探求し、それを提案する。
- ・ 体系化された解釈説明の要素の中で、現状では欠落または不十分な要素を明確にして、それに対する有望なアプローチを探求し、それを提案する。
- ・ 以上の提案検証のために多様なユースケースを収集する。

2-1 研究調査の動機、情報通信との関係、これまでの経緯

【研究調査の動機】

そもそもなぜ仮説の解釈が必要なのであるのか。データ分析応用（AI・機械学習やビッグデータ応用システム）の当事者は、分析者、分野専門家、およびエンドユーザである。一言でいえば、データ分析応用の当事者に仮説を受容してもらうためである。

つまり分析者や専門家は、仮説が中心的な役割を果たす分析応用全体が純粋に技術的に信頼に足るものかどうか判断する必要がある。さらに、サービスの提供者はそのような応用をサービスの享受者に理解させて、利用してもらうためにその説明責任を果たす必要がある。言い換えればサービスの享受者は、サービスの個

別の決定を含む応用全般に関して説明を受ける権利 (right to explanation) を有する[参考文献4] (項目11参照)。そのためには基本として仮説を解釈することが必須となってくる。

例えば科学分野の応用では、分析者と専門家はデータエンジニアと科学者からなるグループである。また社会基盤的応用では、分析者はデータエンジニアであり、専門家は意思決定者であり、エンドユーザはサービスを受ける一般人である場合が多い。

【情報通信との関係】

実用化に向けて研究開発が進む情報通信として説明可能なAIや解釈可能な機械学習の研究([参考文献5, 6, 7, 8]など)は活発である。しかしその大半は個別的な技術の開発が中心であり、解釈説明の要素とその関係からなる解釈説明の全体的構造が明らかではないために、その中で位置付けや体系化が十分になされていない。さらに現状の仮説生成の記述にはプログラミング言語が用いられ、記述のレベルが低いために、仮説の解釈説明に適する手段の探求と提案が必要である。さらに科学的研究手法を研究の対象とする科学哲学においても解釈説明の研究[参考文献9]は行われてきたが、科学哲学の成果と現状の情報通信技術は十分には関連付けられていない。

【これまでの経緯】

東京都福祉保健局と協働でCOVID-19の新規陽性者数予測モデルの構築[参考文献15]やJAXAと協働した、月周回衛星かぐやデータを用いた月探査の研究[参考文献14]によって事業担当者や科学者から仮説・モデルの構築方法や特徴、判断理由の説明をしばしば求められた。こうした経験から解釈説明が仮説を受け入れるためには、必須であり、特に医療や科学の分野では仮説の精度より重視されることを痛感した。

2-3 研究調査の方法・進め方

以下のような手順で研究を行う。

- ・ ユースケースの収集と分析を行う。

ここでは地域観光、スマートシティ、月惑星科学などできるだけ異なる分野からユースケースを収集する点で従来の応用研究とは異なる。

- ・ 解釈説明の構造と種類の明確化

仮説解釈の構造とその構成要素となる説明の種類を明らかにする。従来の研究ではグローバルな説明とローカルな説明という分類を行う研究は存在したが、より深く解釈構造の探求に焦点を当てた研究はなかった。

- ・ 個別技術の体系化

提案する解釈説明の構造に基づいて個別の技術の体系化を図る。従来は解釈説明の構造が無かったので、十分な体系化がなされてこなかった。

- ・ 解釈説明の記述手段の検討

仮説の解釈説明に適した仮説の記述手段を探求し、それを提案する。従来はpythonに代表されるプログラミング言語が仮説生成の手段として使われてきたため、仮説の解釈説明には低水準という問題点であった。

- ・ 今後研究すべき方向性の明確化

体系化された解釈説明の要素の中で、現状では欠落または不十分な要素を明確にして、それに対する有望なアプローチを探求し、それを提案する。従来の解釈説明の構造が明確でなかったために、補うべき新たな技術の明確化が行われにくかった。また科学哲学における解釈説明の研究成果が十分に活かされていなかった。

全体的な検討を代表者が中心となってい、ユースケースの収集と検証に必要となる技術の実装と検証を分担者が中心となって進める。

2-4 研究調査に関する内外の研究の動向

説明可能な人工知能 (eXplainable AI, XAI) や解釈可能な機械学習 (Interpretable Machine Learning, IML) の分野でマイクロな説明機能に相当する研究が活発化し始めた[参考文献5]。XAI や IML に関する個別研究は (1) 特徴量に基づく説明, および (2) 解釈可能なモデル, (3) 説明モデルの導出に分類される。 (1) には説明ベクトルの導入や弱分類器の利用により重要な属性を発見する研究 [参考文献6] や [参考文献13] がある。 (2) には文字の生成過程を考慮することで解釈できるモデルを作る研究 [参考文献7] がある。 (3) には任意の分類器から, 結果に局所的な近似モデルを用いてその説明を導出する研究 [参考文献8] や可視化する研究 [参考文献12] がある。すでに解釈説明の概念化の検討 [参考文献14] には着手したが, 本研究では, 現在はまだ十分に研究されていない解釈説明構造を明らかにする。次に本研究はマイクロな説明技術として, 生成モデルの説明と個別決定の理由の説明にそれぞれ上記 (1) と (3) カテゴリのアプローチを発展させつつ, 従来十分に研究されていないマクロな説明技術の研究を推進し, 両者を統合発展させた画期的な解釈説明基盤の構築を目指す。その際に構造に欠落する要素技術を探求する。さらに生成モデルに依存しない高水準の言語としてデータベース操作言語 SQL [参考文献10] が検討対象になる。

3 研究調査の成果と今後の計画

3.1 研究調査の成果

- ・ ユースケースの収集と分析を行う。

分析対象として, 地域観光 (発表資料2, 3, 4), スマートシティ, 月惑星科学, ソフトウェア品質管理 (発表資料1) などの異なる分野からユースケースを収集した。

- ・ 解釈説明の構造と種類の明確化

仮説解釈の構造とその構成要素となる説明の種類を明らかにした (発表資料5, 6)

説明の対象としては, 以下のような構成要素を仮定して, 解釈説明の構造を提案した (図1)。

- ・ データの作り方 (how_d) : How to generate Data
- ・ 仮説の作り方 (how_h) : How to generate Hypothesis
- ・ 仮説の特徴 (what_h) : What features of Hypothesis
- ・ 仮説結果の理由 (why_h) : Why (What Reasons) for Hypothesis decision
- ・ 生成方法の実行計画 (how_e) : How to Execute data/hypothesis

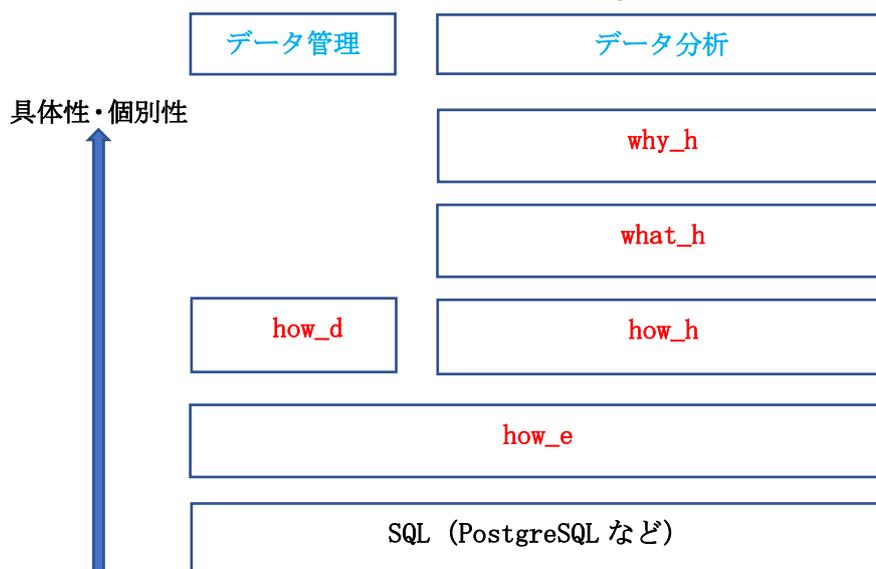


図1

提案する要素は、具体性・個別性の軸とデータ管理・データ分析というエコシステムの軸の二つで構造化されている。

従来の研究ではグローバルな説明とローカルな説明という単純な分類を行う研究は存在したが、このような深い解釈構造を導入することによって、これまで以上に解釈構造の探求分類に焦点を当てることができるようになった。

- ・ 個別技術の体系化

従来は解釈説明の構造が無かったので、十分な体系化がなされてこなかったが、提案する解釈構造に基づいて、より包括的に個別の技術の体系化を図ることができ、その結果を著書にまとめた（発表資料 5, 6）。

- ・ 解釈説明の記述手段の検討

従来は python に代表されるプログラミング言語が仮説生成の手段として使われてきたため、仮説の解釈説明には低水準であるという問題点があった。仮説の解釈説明に適した仮説の記述手段として、SQL などの非手続き的言語を中心に探求し、それを提案した（発表資料 6）。

- ・ 今後研究すべき方向性の明確化

従来は解釈説明の構造が明確でなかったために、補うべき新たな技術の明確化が行われにくかった。

まず、隣接する分野のひとつである科学哲学における解釈説明の研究成果が十分に活かされているとは言えなかったことである。そこで今後は科学哲学を含め、異分野の解釈説明の成果を生かすべきである。

さらに体系化された解釈説明の要素の中で、現状では欠落または不十分な要素を明確にできていなかった。そうした要素の中には、まだ十分に定義されていない解釈説明を生成するための一般的なアーキテクチャが含まれる。そこで、それらに対する有望なアプローチの検討を開始した（発表資料 6）。

3.2 研究調査の将来計画

本研究期間の終了後は、構築した統合的仮説生成・説明方式と実装アルゴリズムを複数のユースケース（地域観光やスマートシティ、月惑星科学など）に適用して、その妥当性の検証を行い、ソーシャルビッグデータの利活用と流通を促進すると同時に情報通信分野において、言わば「解釈説明工学」のような新領域の構築を目指す。

将来的には、本研究の成果により、ビッグデータを媒介として、これまで縁遠かった異業種セクタが協業できるようなダイナミックな知的社会 Society5.0 を実現することが期待できる。その意味で、本研究は社会的意義も十分に高いと考えられる。

具体的には、以下のように成果を公表するとともに、研究終了後も研究成果を発展させる計画である。

- ・ 著作や論文発表を通して、研究成果を公表する。
- ・ 国際会議・研究フォーラムにおいて仮説解釈説明に関連するスペシャルトラックなどを開催して研究成果の発表をおこなうと同時に、関連する研究者の交流する機会を設けて、この分野の強化を行う。
- ・ 所属する大学において国際研究拠点の構築を推進する。
- ・ HPで研究成果を周知するとともに所属する大学において社会人向け講座を開催し、授業を行う。

さらに新しいユースケースへの適用による提案の妥当性を確認していくとともに、新領域の確立を目指して、新たに発掘された研究課題に取り組んでいく。

【参考文献】

1. 内閣府, Society 5.0. https://www8.cao.go.jp/cstp/society5_0/index.html
2. NRI, 「情報銀行」は消費者に受け入れられるか?. NRI Journal, 2018.
3. EU GDPR. <https://gdpr.eu/>

4. M.E.Kaminski, The Right to Explanation, Explained. University of Colorado Law School, 2019.
5. C.Molnar, Interpretable Machine Learning, <https://christophm.github.io/interpretable-ml-book>
6. D.Baehrens et al., How to explain individual classification decisions. JMLR, 11, 2010.
7. B.M.Lake et al., Human-level concept learning through probabilistic program induction. Science 350(6266):1332-1338, 2015.
8. M.T.Ribeiro et al.,“Why Should I Trust You?” Explaining the Predictions of Any Classifier. <https://arxiv.org/abs/1602.04938>
9. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/>
10. U.Malik et al., SQL for Data Analytics, Packt. 2019.
11. 石川博, 仮説のつくりかた : 多様なデータから新たな発想をつかめ. 320 頁, 共立出版, 2021.
12. Hiroshi Ishikawa, Yasushi Miyata, Social Big Data: Case Studies. Trans. Large Scale Data Knowl. Centered Syst. 47, pp. 80-111, Springer Nature, 2021.
13. Hiroshi Ishikawa, Yukio Yamamoto, Masaharu Hirota, Masaki Endo, An Explanation Framework for Whole Processes of Data Analysis Applications: Concepts and Use Cases. Intl Journal on Advances in Software, vol.13, No.1&2, pp.1-15, 2020.
14. 加藤広大, 山田竜平, 山本幸生, 廣田雅春, 横山昌平, 石川博, 惑星の位置関係に基づく深発月震分類のための特徴量の検討. Journal of Space Science Informatics Japan, Vol.7, pp.43-52. 2018.
15. Yusuke Takamori, Riku Watanabe, Daiju Kato, Masaki Endo, Hiroshi Ishikawa: Numerical estimation of new COVID-19 positive cases using time series analysis by machine learning. Proc. Intl. Conf. ACM MEDES, 2021.

〈 発 表 資 料 〉

題 名	掲載誌・学会名等	発表年月
[1] Quality Control Methods Using Quality Characteristics in Development and Operations.	Digital 2024, 4	2024.3
[2] Time-Series Estimation of the Best Time to View Seasonal Organisms Using Geotagged Tweets and Co-occurring Words.	Communications in Computer and Information Science, vol 2022. Springer Nature	2024.2
[3] Estimating the best time to view cherry blossoms using Natural Language Processing tasks	INFSOC IWIN 2023	2023.9
[4] Relation between time-series forecasting methods and estimation of best times to see cherry blossoms	INFSOC IWIN 2023	2023.9

[5] Hypothesis Generation and Interpretation: Design Principles and Patterns for Big Data Applications	Springer Nature ISBN 9783031435393	2024.2
[6]モダン SQL データ管理から分析へ	共立出版 ISBN 9784320125698	2023.11