

最大 73 倍高速化など、あらゆる形式のデータ分析に対して 世界最速級の情報選択 AI アルゴリズム群「高速スパースモデリング技術」を確立

～製造・医療・マーケティングなど、最大 1 ヶ月かかる分析を 1 日以下に短縮し、
大量の遊休データを分析可能に～

発表のポイント:

- ◆ 大量のデータから重要な情報を選択するデータ分析を、従来アルゴリズムと比較して同様の精度を保ちつつ最大 73 倍高速化する AI アルゴリズム群「高速スパースモデリング技術」を確立しました。
- ◆ 製造・医療・マーケティング・エネルギーなどの様々な分野で取得される多様なデータ形式に対応しており、あらゆる業界の分析待ちの解消を期待できます。
- ◆ 本アルゴリズム群の一部は NTT ドコモビジネスのノーコード AI 開発ツール「Node-AI」にて利用可能です。

NTT 株式会社(本社:東京都千代田区、代表取締役社長:島田 明、以下「NTT」)は、大量のデータから重要な情報を選択するデータ分析を従来アルゴリズムと比較して最大 73 倍高速化する世界最速級の情報選択 AI アルゴリズム群「高速スパースモデリング技術」を確立しました。本技術は従来のスパースモデリング技術と同じ分析精度となることを理論的に保証していることから、質を落とすことなくデータ駆動型の意思決定を迅速化できます。また、本アルゴリズム群は多様なデータ形式に適用できるため、製造・医療・マーケティング・エネルギーなどのあらゆる業界で生じている大量データの分析待ちを解消できると期待されます。本アルゴリズム群の一部は、NTT ドコモビジネスのノーコード AI 開発ツール「Node-AI」に既に搭載されて利用可能です。本成果の一部は、ECML PKDD(*1) 2025 において発表される他、NeurIPS(*2) 2019 および 2024、ICML(*3) 2020、AISTATS(*4) 2023 に掲載されました。

1. 背景

近年、センシング技術の進展により、ヒトやモノから大量の多種多様なデータを取得できるようになりました。NTT ではこうしたデータを分析することにより様々な分野でデータ駆動型の意思決定を行っています。その中でも大量のデータから重要な情報を選択する分析は結果の解釈性が高く、その後の分析方針の立案や戦略の仮説を作成するために分析の初期段階でよく実施されます。

スパースモデリングはこのような情報を選択する分析に広く用いられる学習型の技術です。スパ

ースモデリングは「得られた情報の中でも必要なものはごく一部で、その他の大部分は不要である」というスパース性を仮定することにより、大量のデータから重要な一部の情報を選択することができます。

しかし、データが膨大になるほどスパースモデリングの処理時間も長くなり、現実的な時間内でデータを分析することは難しくなります。こうしたデータ分析の長時間化が、週例会議に分析結果を提示できないなどの PDCA サイクルの遅れを引き起こし、最終的な意思決定に大きな後れをもたらしていました。また、分析時間を短縮するためにデータの一部のみを使用するといった対策がとられることもあります。しかし、こうした対策では使われないデータは分析されない遊休データとなってしまうため、重要な知見を見逃してしまうリスクもありました。

2. 成果の概要

①枝刈りによる高速化アルゴリズムの確立： NTT ではスパースモデリングで大量のデータを現実的な時間内で分析するため、その高速化アルゴリズムを開発してきました。スパースモデリングを用いてデータから重要な情報を選択する際は、「必要な情報はごく一部である」と仮定することでこれを実現します。言い換えると、この仮定の元では「不要な情報が大量に存在する」とも解釈できます。この不要な情報に対応する計算を安全にスキップする枝刈りアルゴリズムを開発することで、元のアルゴリズムと比較して理論的にも実験的にも精度を損なうことなく最大 73 倍の高速化に成功し AI・機械学習分野の最難関国際会議 NeurIPS*5*10 や ICML*6 に論文が採択されています(図1)。これにより、例えば 1 か月以上かかるような分析を 1 日以下に短縮し意思決定を迅速化できます。また、これまで時間的制約で分析できなかった遊休データも分析可能となります。

②多様なデータ形式への拡張： 分析の精度を高めるためには、データだけでなくそのデータ形式を付加情報として活用することも重要です。例えば通信網の通信データはネットワーク構造形式、地域別交通量データはグループ構造形式、EC サイトの商品データは商品分類の階層(木)構造形式を有しています。データと共にこれらの構造形式も情報として分析に活用することで分析精度の向上を期待できます。NTT は上述の枝刈りアルゴリズムをグループ構造形式*7 やネットワーク構造形式*8 へ適用することに成功し、AI・機械学習分野の難関国際会議 AISTATS*4 などに論文が採択されています。今回、階層(木)構造形式*9 に対応した枝刈りアルゴリズムが機械学習・データマイニング分野の難関国際会議 ECML PKDD*1 に採択されました。これにより本アルゴリズム群が実用的なデータ形式の大半をカバーできるようになり(図 1)、あらゆる業界の分析待ちの解消を期待できます。

	データ形式	主要適用分野	設定	技術名	最大高速化倍率	従来技術	論文採択先	製品導入
①枝刈りによる高速化技術の確立	無	マーケ・エネルギー	教師あり	FastIHT	73倍	IHT	NeurIPS'24	-
			教師なし	FastCUR	10倍	CUR	ICML'20	-
②多様なデータ形式への拡張	バイレベル	製造	教師あり	FastSGL	35倍	SGL	NeurIPS'19	ドコモビジネス Node-AI
	グループ	医療・ヘルスケア	教師あり	FastGSCAD	68倍	GSCAD	AISTATS'23	ドコモビジネス Node-AI
	グラフ	通信・交通	教師あり	FastGraph Lasso	3倍	Graph Lasso	JSAT'21	-
	木	EC・気象	教師あり	FastTreeLasso	2倍	TreeLasso	ECML PKDD'25	-

図 1: 本技術が対応するデータ構造・高速化倍率・論文採択先・利用状況

3. 技術のポイント

①スパースモデリングによる分析を最大 73 倍高速化： スパースモデリングにおける、「必要な情報はごく一部である」というスパース性の仮定を活かし、不要な情報に対応する計算を安全にスキップして枝刈りする独自のアルゴリズムを導入することで、元のアルゴリズムと比較して最大 73 倍の高速化に成功しました(図 2)。

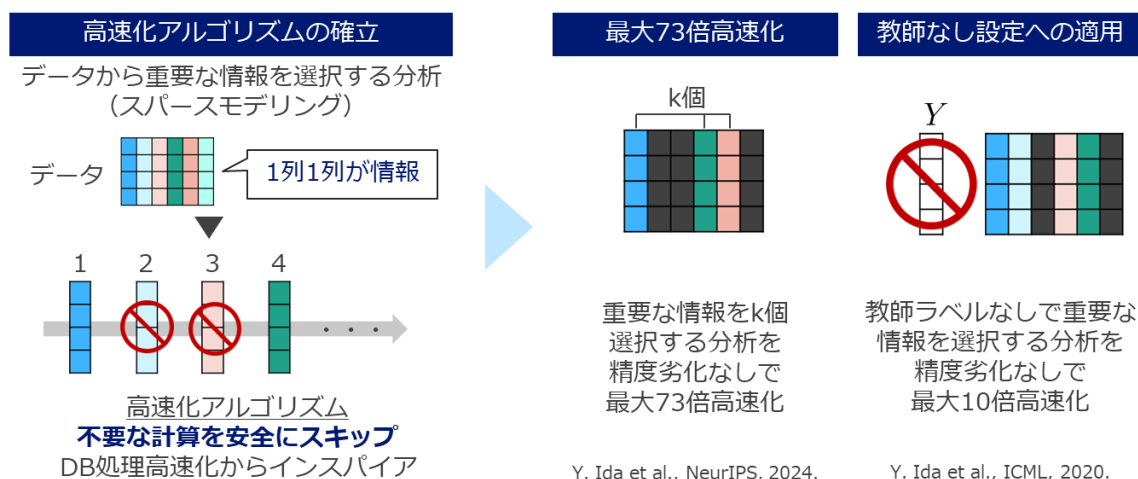


図 2: 枝刈りによるスパースモデリング高速化アルゴリズムの確立

②従来技術から精度劣化しないことを理論保証： 本技術は従来技術における不要な計算のみをスキップして枝刈りするため、従来技術と比較して精度が劣化しないことを理論的に保証可能です(図 3)。

- ▶ スパースモデリングは多くの情報から重要な一部の情報を選択する技術。
 - ▶ 各情報の重要度を計算し、それが**閾値**以下であれば**その情報は重要でない**と判定。
- ▶ 高速スパースモデリングは重要度の代わりに、**高速に計算できる重要度の上界**を使って判定。

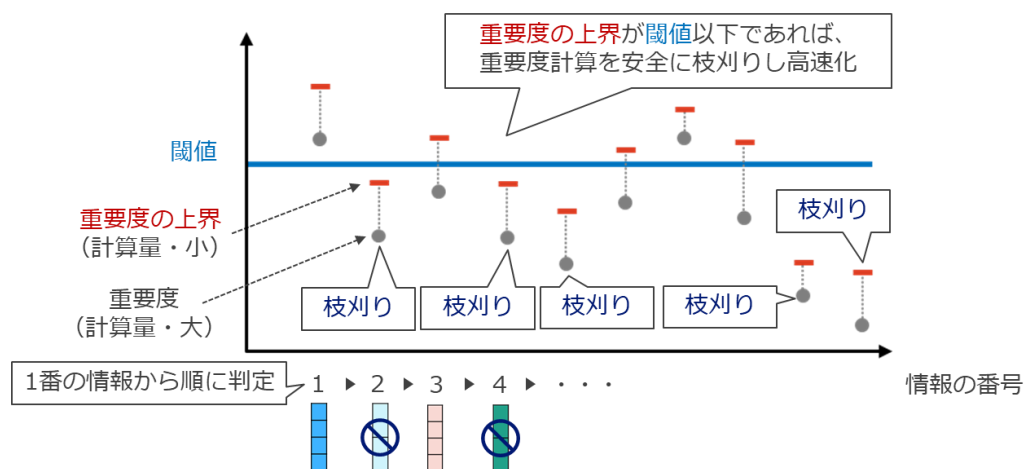


図 3: 本技術の枝刈りイメージ

③実用的なデータ形式を網羅： 本技術はデータがグループ構造、ネットワーク構造、階層（木）構造などを有している場合にもこれらの情報を加味して分析可能であり、データ自体に加えてこのような構造情報を利用することで、より様々なドメイン特化型の分析精度の向上を期待できます（図4）。

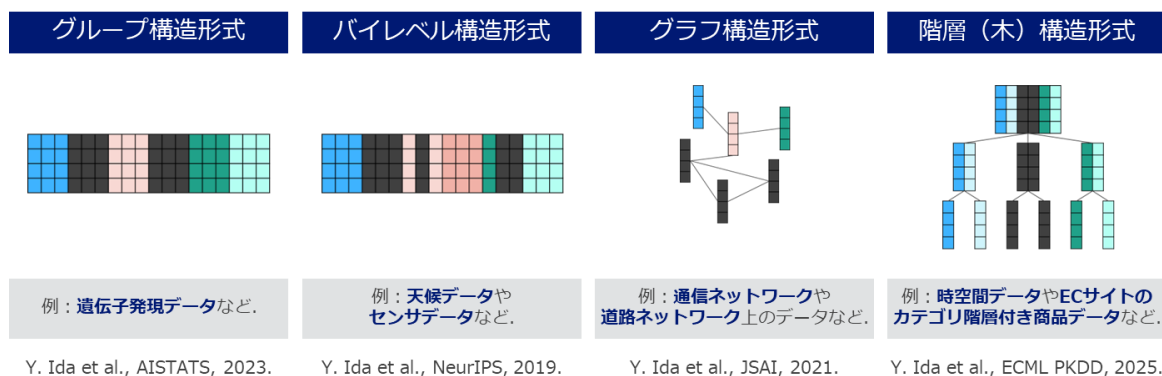


図 4: 本技術が対応するデータ構造形式のイメージとデータ例

4. 応用例

製造分野におけるプラントの生産効率最適化において、プラントに設置されたセンサから得られる時系列データから、プラントの生産の増減要因を推定します。本技術により生産量に影響した時刻を時系列データの中から特定できれば、その時刻でどのような操作が行われていたか、などの状況を調べる糸口となり、より生産効率を高める制御操作の検討や策定を期待できます。ここで、センサが1秒毎にデータを取得していた場合、1日で約1万時刻分のデータを取得できます。取得期間が長くなるほどデータ量は増加し、数万～数億時刻となると一般的なコンピュータで1ヶ月以上の分析時間を要する場合があります。データの性質や問題設定に依存しますが、本技術は上記の分析時間を1日以下へ短縮できるケースもあります。

その他にも医療分野において特定の病気の遺伝的要因の推定を行い新薬や治療法開発にその分析結果を役立てる、マーケティング分野において顧客の購買行動から特徴的な行動を推定し広告やクーポンの最適化に役立てる、エネルギー分野において核融合炉のセンサデータからプラズマの挙動を表す方程式を推定し核融合炉の安定稼働のための制御操作策定に役立てる、などの応用例が考えられます（図5）。

- ▶ 高速スパースモデリング技術は製造・医療・マーケティング・エネルギーなどの様々な分野のデータ分析で活用可能。

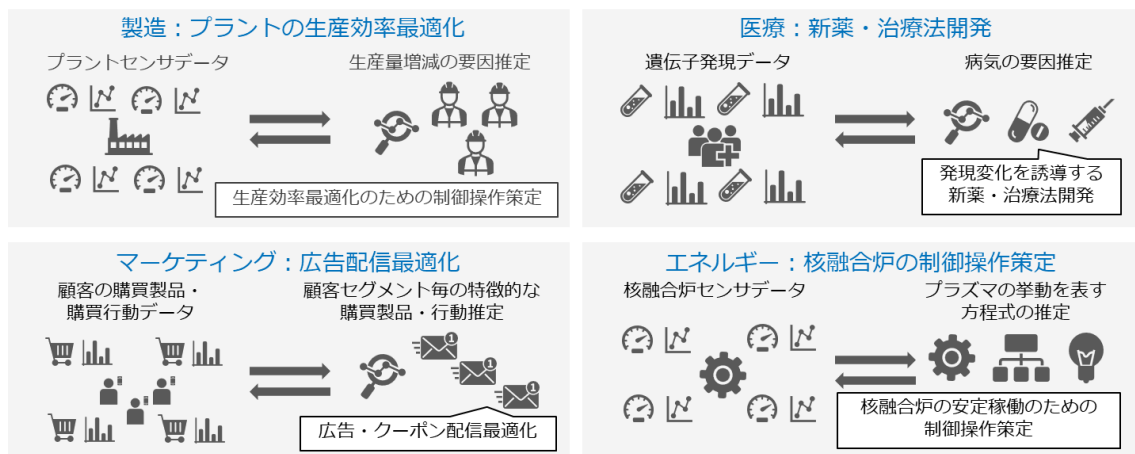


図 5: 本技術の分野毎の応用例

5. 製品搭載の概要

本高速スパースモデリング技術の一部は既に NTT ドコモビジネスのノーコード AI 開発ツール「Node-AI」*11 に搭載されており利用が可能です。「Node-AI」はコードを書かずにデータ分析 AI を構築できるツールです。時系列データの分析に強く、AI の予測の根拠を出力する要因分析などをサポートしています。高速スパースモデリング技術の搭載により、コードを書くことなく高速かつ高精度な情報選択が可能となります※1※2。

6. 今後の展開

本技術は多様なデータ形式をサポートすることで様々な分野に適用可能であるため、「Node-AI」への搭載をはじめとした実用化を推進し、NTT グループのビジネス領域拡大に貢献します。また、引き続き高精度化や省メモリ化といったスパースモデリングの高度化に向けた研究開発を推進していくとともに、大量データの分析が不可欠な核融合等の自然科学分野への適用(AI for Science*12)など、多様な応用に取り組んで参ります。

【用語解説】

*1 European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases。機械学習・データマイニングに関する難関国際会議。

*2 Annual Conference on Neural Information Processing Systems。AI・機械学習に関するトップレベルの最難関国際会議。

*3 International Conference on Machine Learning。AI・機械学習に関するトップレベルの最難関国際会議。

*4 International Conference on Artificial Intelligence and Statistics。AI・機械学習に関する難関国際会議。

*5 Y. Ida, S. Kanai, A. Kumagai, T. Iwata, and Y. Fujiwara, “Fast Iterative Hard Thresholding Methods with Pruning Gradient Computations”, Advances in Neural Information Processing Systems (NeurIPS), 2024.



*6 Y. Ida, S. Kanai, Y. Fujiwara, T. Iwata, K. Takeuchi, and H. Kashima, “Fast Deterministic CUR Matrix Decomposition with Accuracy Assurance”, International Conference on Machine Learning (ICML), 2020.

*7 Y. Ida, S. Kanai, and A. Kumagai, “Fast Block Coordinate Descent for Non-Convex Group Regularizations”, International Conference on Artificial Intelligence and Statistics (AISTATS), 2023.

*8 Y. Ida, Y. Fujiwara, and H. Kashima, “Fast Block Coordinate Descent for Sparse Group Lasso”, The Japanese Society for Artificial Intelligence, 2021.

*9 Y. Ida, S. Kanai, A. Kumagai, T. Iwata, and Y. Fujiwara, “Fast Proximal Gradient Methods with Node Pruning for Tree-Structured Sparse Regularization”, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD), 2025.

*10 Y. Ida, Y. Fujiwara, and H. Kashima, “Fast Sparse Group Lasso”, Advances in Neural Information Processing Systems (NeurIPS), 2019.

*11 Node-AI, <https://nodeai.io/>

*12 科学研究のプロセス全体を AI で加速させる取り組み。

※1 高速で解釈性の高い FastSQL モデル (Ver.3.9.6),
<https://note.com/nodeai/n/nb947c9e692ac>

※2 学習速度と変数選択精度を極めた FastGSCAD モデル (Ver. 3.22.0),
<https://note.com/nodeai/n/nf879b51f36bf>

■本件に関する報道機関からのお問い合わせ先

NTT 株式会社

サービスイノベーション総合研究所

広報担当

[問い合わせフォームへ](#)