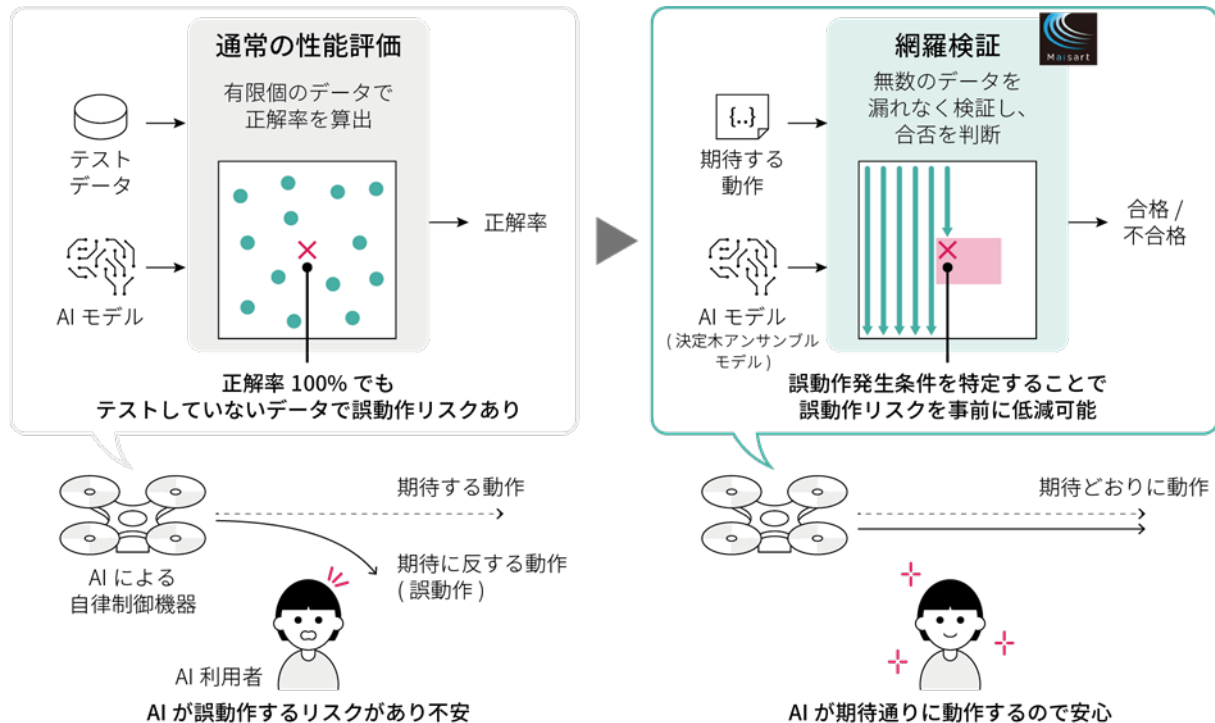


NEWS RELEASE

AI の信頼性向上に寄与する「AI の動作を短時間で漏れなく検証する技術」を開発
決定木アンサンブルモデルの誤動作リスクを低減し、安心して AI を利用できる社会の実現に貢献



AI の通常のパフォーマンス評価と網羅検証の比較

三菱電機株式会社は、決定木アンサンブルモデル※1を対象とした「AI の動作を短時間で漏れなく検証する技術」を開発しました。本技術は当社 AI 技術「Maisart® (マイサート) ※2」の開発成果で、AI が期待に反する動作（誤動作）を行うリスクを低減し、安心して AI を利用できる社会の実現に貢献します。

AI が急速に発展して世界的に普及が進む中、AI に関する世界初の包括的な法的枠組みを定めた「欧州 AI 規制法」が 2024 年 8 月 1 日に発効されました。また各国でも同様に、AI のリスクに対処するための法令やガイドラインの整備が進められており、AI を開発・提供する事業者には、リスクを適切に管理することが求められています。特に、機器の自律制御システム、電力・社会インフラシステム、サイバーセキュリティシステムなどの安全性が重視されるシステムにおいては、誤動作時に発生する損害の大きさから AI の信頼性が重要となっています。通常、AI の信頼性評価は、AI モデルの学習に用いていない有限個のテストデータで正解率などの指標を評価しますが、AI の動作は非常に複雑で、仮にテストを実施した際の正解率が 100% であっても、テストしていないデータでの誤動作リスクを排除できないという問題がありました。この問題に対して、AI に期待する動作をあらかじめ設定し、期待どおりに動作しているかを厳密に漏れなく検証する網羅検証の手法が提案されています。ただし、AI モデルのサイズが大きいと検証に膨大な時間がかかることや、誤動作リスクの大きさが不明で対処の優先度を判断しづらいことなどを理由に、網羅検証を実施するケースは限定的となっていました。

当社は今回、網羅検証における問題を解決する技術として、数値データの予測などに広く用いられる決定木アンサンブルモデルに対して効率的に網羅検証を行う新たなアルゴリズムと、これを用いた対話的な検証ツールを開発しました。本技術を活用することで、AI 開発者は網羅検証のサイクルを高速で循環させることができるため、AI の誤動作リスクを低減することが可能となります。これにより、AI の信頼性を向上し、安心して AI を利用できる社会の実現に貢献します。

※1 データを条件に基づいて分割して予測を行うモデルである決定木を複数組み合わせる予測精度を向上させる AI 手法

※2 Mitsubishi Electric's AI creates the State-of-the-ART in technology の略。
全ての機器をより賢くすることを目指した当社の AI 技術ブランド

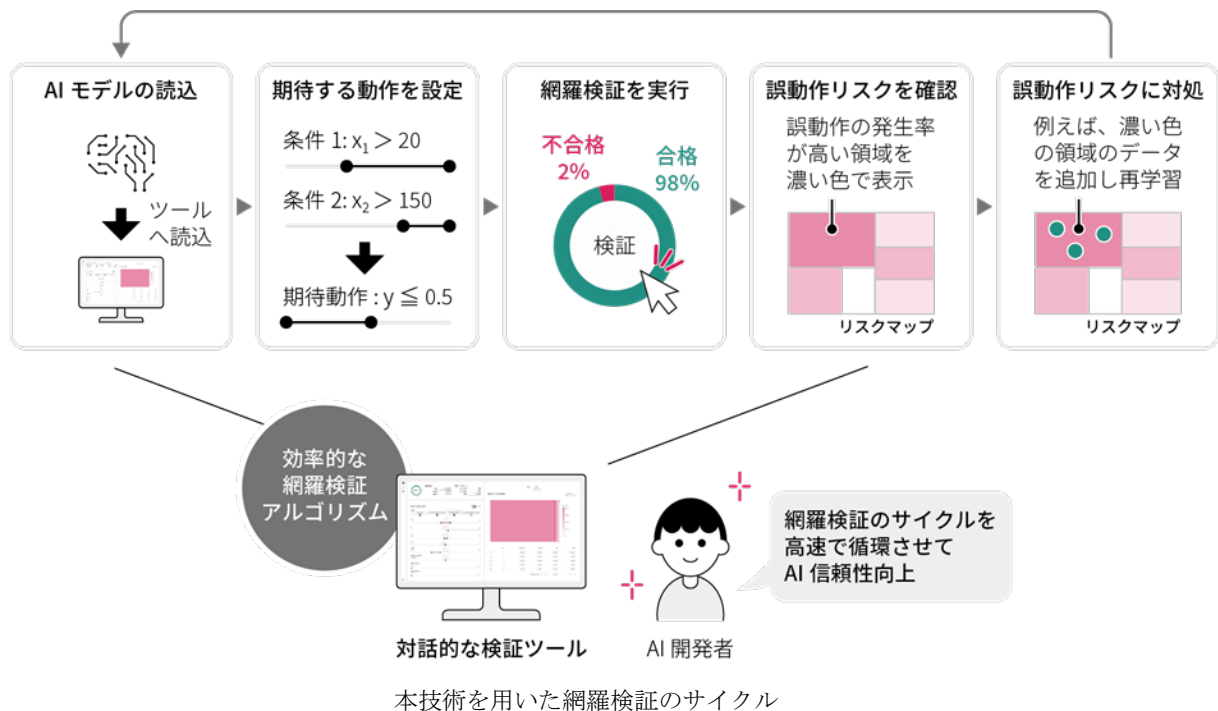
開発の特長

1. 効率的に網羅検証を行うアルゴリズムの開発により、AI の誤動作リスクを短時間で検証可能

- ・決定木アンサンブルモデルに対して、効率的に網羅検証を行うアルゴリズムを開発。これにより、AI が期待どおりに動作しているかを厳密に漏れなく検証可能
- ・AI に入力するデータの範囲を再帰的に分割して検証することで、網羅検証における従来の手法と比べて数十～数百倍高速^{※3}に検証できることを確認

2. 直感的に操作できる対話的な検証ツールを開発、誤動作リスクへの適切な対処を支援

- ・ブラウザベースの GUI^{※4}を介して、ユーザーが直感的な操作で網羅検証を実施できるツールを開発。学習済みの AI モデルを読み込み、AI に期待する動作を設定することで、ワンクリックで検証を実行
- ・検証結果として合否の割合を表示することで、AI 開発者が誤動作リスクの大きさを把握可能
- ・誤動作の発生条件ごとの発生率を色の濃淡で表したリスクマップとして図示することで、AI 開発者の誤動作リスクへの適切な対処を支援



今後の予定・将来展望

2025 年度以降、社内外で開発する AI を対象に実証を進めていきます。また AI の標準化活動とも連携し、開発技術を広く社会に還元していくことで、安心して AI を利用できる社会の実現を目指します。

商標関連

「Maisart」

三菱電機株式会社の登録商標

参考情報

効率的な網羅検証アルゴリズムの詳細はデータマイニング分野の国際会議 IEEE ICDM 2024 のワークショップ^{※5}および人工知能学会全国大会^{※6}の論文を参照ください。

※3 2025 年 2 月 26 日時点、当社調べ

※4 Graphical User Interface

※5 Saori Matsunaga, and Genta Yoshimura “Efficient and High-Quality Formal Verification for Decision Tree Ensembles” International Workshop on Adaptable, Reliable, and Responsible Learning (ARRL2024)

※6 松永 沙織、吉村 玄太「決定木アンサンブルモデルの形式検証と検証結果の要約」人工知能学会全国大会 (JSAI2024)

三菱電機グループについて

私たち三菱電機グループは、たゆまぬ技術革新と限りない創造力により、活力とゆとりある社会の実現に貢献します。社会・環境を豊かにしながら事業を発展させる「トレード・オン」の活動を加速させ、サステナビリティを実現します。また、デジタル基盤「Serendie®」を活用し、お客様から得られたデータをデジタル空間に集約・分析するとともに、グループ内が強くつながり知恵を出し合うことで、新たな価値を生み出し社会課題の解決に貢献する「循環型 デジタル・エンジニアリング」を推進しています。1921 年の創業以来、100 年を超える歴史を有し、社会システム、電力システム、防衛・宇宙システム、FA システム、自動車機器、ビルシステム、空調・家電、情報システム・サービス、半導体・デバイスといった事業を展開しています。世界に 200 以上のグループ会社と約 15 万人の従業員を擁し、2023 年度の連結売上高は 5 兆 2,579 億円でした。詳細は、www.MitsubishiElectric.co.jp をご覧ください。

お問い合わせ先

< 報道関係からのお問い合わせ先 >

三菱電機株式会社 広報部

〒100-8310 東京都千代田区丸の内二丁目 7 番 3 号

TEL 03-3218-2332

<https://www.MitsubishiElectric.co.jp/news/contact.html>

< お客様からのお問い合わせ先 >

三菱電機株式会社 情報技術総合研究所

〒247-8501 神奈川県鎌倉市大船五丁目 1 番 1 号

https://www.MitsubishiElectric.co.jp/corporate/randd/inquiry/index_it.html