

AIを適用したIIoTサービスの継続的運用を可能にするMLOps基盤

MLOps Platform for Continuous Operation of AI-Based IIoT Services

小島 知也 KOBATAKE Tomoya 山田 正隆 YAMADA Masataka

様々な要因によってデータが変化する実サービスにおいてAIを運用するには、継続的にAIモデルを更新することで変化に追従していく必要がある。しかし、AIモデルの更新は、モニタリングや再学習、AIモデル管理といった専門性の高いプロセスから成るため、多大な人的リソースを必要としていた。

東芝グループは、これらのプロセスを自動化して現場での円滑な実行を可能にする、MLOps (Machine Learning Operations) 基盤を開発し、AIを適用したIIoT (Industrial Internet of Things) サービスの開発速度・品質の向上を図っている。このMLOps基盤を、プラントの異常予兆検知や製品の自動検査のためのIIoTサービスなどに導入し、AIの専門家だけでなくAIモデルの更新プロセスを実行してデータ変化に追従させることが可能であることを確認した。

In order to maintain the performance of artificial intelligence (AI) applied to the provision of actual services, where data are subject to change due to various factors, it is necessary to continuously update the AI model. However, the updating of an AI model consists of highly specialized processes such as monitoring, relearning, and AI model management, requiring substantial human resources.

The Toshiba Group has developed a machine learning operations (MLOps) platform that automates these processes and allows them to be smoothly implemented on site, thereby improving the speed and quality of development of AI-based Industrial Internet of Things (IIoT) services. We have applied this MLOps platform to IIoT services such as plant anomaly prediction and automatic product inspection, and confirmed that it is possible for personnel who are not AI specialists to execute the process of updating AI models and to take the necessary measures in response to changes in data.

1. まえがき

近年、AIの成長は目覚ましく、様々な分野において優れた性能を発揮している⁽¹⁾。東芝グループが力を入れているIIoTサービスの分野も例外ではなく⁽²⁾、インフラ設備の予知保全や製品検査工程の自動化など、AIが適用可能な範囲も多岐にわたる。東芝グループは、AIのアルゴリズムに関する研究開発を活発に行っており⁽³⁾、世界でも有数の特許数を誇っている⁽⁴⁾。

しかしながら、これらのAIを実サービスに適用し、継続的に運用するには、様々な問題がある⁽⁵⁾。その一つに、AIモデルの性能は学習に使用するデータに依存しており、推論時のデータの傾向が学習時と異なる場合、所望の性能が得られないことが挙げられる。このようなAIの特性は、季節変動や設備の劣化・改修などの影響によりデータが変化する実サービスにおいてAIを活用する際に、大きな障壁となる。特に、データ変化の要因や傾向を事前に予測できない場合には、継続的に再学習を行うことでAIモデルを更新し、データ変化に追従していくことが必要になる。

しかし、AIモデルを継続的に更新するには、データや推

論精度のモニタリングをはじめ、データの収集・選別や、学習パラメーターのチューニングによる再学習、性能比較・更新判断といったAIモデル管理など、専門性の高い複数のプロセスを繰り返し実施する必要があり、AIの活用現場で運用するにはハードルが高い。このようなAI運用時の課題解決に重点を置き、継続的に再学習を行う手法はMLOpsと呼ばれ、AIを実サービスに適用する事例が増えてきたことで注目されている⁽⁶⁾。

そこで東芝グループは、様々なIIoTサービスを手掛けている経験を生かし、複数の現場における課題を基に前述したプロセスの自動化を実現するMLOps基盤を開発して、現場適用を行った。これにより、AIを搭載した個々のIIoTサービスにおいては、運用の効率化や、性能の維持・改善が可能になった。更に、共通基盤として様々な事業へ展開することで、インターフェースの統一や運用ノウハウの集約を実現し、東芝グループ全体としてのAIサービス開発・運用の効率化を目指している。ここでは、開発したMLOps基盤の機能構成と適用事例について示す。

2. 機能構成

東芝グループが開発したMLOps基盤は、AIサービスの運用時に繰り返し必要となる一連のプロセスを自動化する(図1)。具体的には、学習を自動化するパイプライン管理機能、AIモデルを推論サービスとして提供するサービング機能、データ変化・AIモデルの性能を監視するモニタリング機能、学習用データを管理する学習データ管理機能、及びソースコードやAIモデルのテスト・適用を自動化するCI/CD(継続的インテグレーション/継続的デリバリー)機能を提供する。これらの機能群を持ったMLOps基盤は、自動構築によって容易に実現できる。

2.1 パイプライン管理機能

パイプライン管理機能は、事前に定義した一連のプロセスをパイプラインとして登録することで、GUI(Graphical User Interface)やAPI(Application Programming Interface)からの自動実行を可能にする機能である。データ取得・データ前処理・学習・評価などのAIの学習プロセスの流れをパイプラインとして定義することで、運用時にはプログラミングの知識がなくても、必要なパラメータを入力するだけでAIの学習を手軽に実行できる。これにより、AIの専門性が要求されるAIモデルの学習プロセスを、現場担当者でも実施可能となった。また、この機能は、パイプライン実行時の入力パラメータ・実行履歴・評価結果を自動で保存し、GUIでアクセス可能としている。AIの学習プロセスでは、所望の精度を達成するため、過去の実行履歴に基づいてパラメータなどを修正しながら繰り返し学習を行う必要があり、自動保存機能が有効である。これにより試行錯誤

の結果の共有が容易になり、複数プロジェクトを遠隔から一元管理できる状態が実現できた。

データ前処理などのパイプライン中で実行される各処理は、コンポーネントとしてコンテナと呼ばれる実行環境を包含した仮想マシンの形で実装される。そのため、個々のコンポーネントは、独立して開発・利用でき、またインターフェースを統一することで、複数のパイプラインにおいて流用できるようになる。これにより、学習部分はAIアルゴリズム開発者が、データ取得部分はITエンジニアが実装するといった役割分担を明確化でき、優れたAIアルゴリズムなどの研究成果物を、コンポーネント単位で様々な事業に利用・展開することが容易になる(図2)。

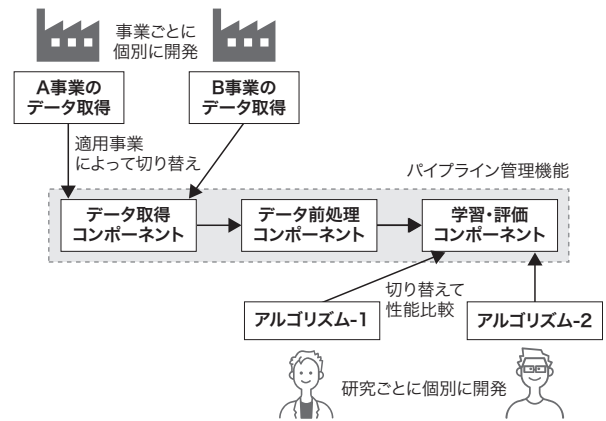


図2. パイプラインとコンポーネント

あらかじめインターフェースを定めてコンポーネント化することで、独立に開発したアルゴリズムなどを切り替えて様々な事業に利用・展開できる。

Pipeline and its components

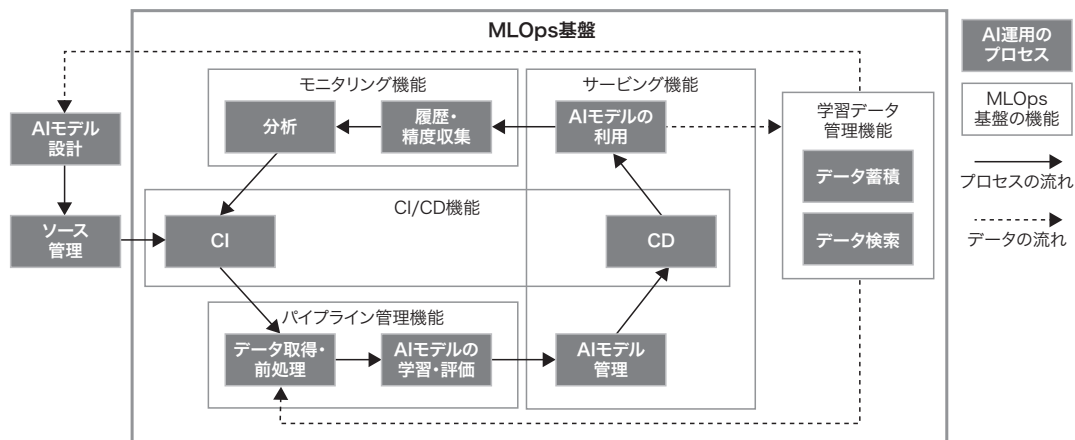


図1. AIサービスの運用プロセスと開発したMLOps基盤の機能群

AIサービスの運用において、従来行われていた専門性の高いプロセスの繰り返しを、MLOps基盤の機能でサポートすることで、効率化を実現する。

AI service operation process and functions of MLOps platform

2.2 サービング機能

学習済みAIモデルを用いた推論サービスをAPIサービスとして提供するサービング機能を導入した。これによりデプロイ(配置)されるAIモデルはコンテナとして管理され、スケールアウト、冗長化、ログ収集、再起動による自動復旧が実現できる。デプロイしたAIモデルの管理を容易にするため、モデルファイルの保存先やデプロイ先、学習に使用したパイプラインなどの情報をメタデータとして基盤上のデータベースに記録し、これらの情報はGUIやAPIを通して簡単に取得できる。

また、開発効率改善のためドメインごとに定義した標準APIサーバーのデプロイを推奨している。これにより推論側のAIアルゴリズムとリクエスト側のアプリケーションを独立して開発でき、アルゴリズムの横展開が容易になる。

2.3 モニタリング機能

MLOpsにおいて重要な役割を果たすのがモニタリングである。実サービスにAIを適用する際には、用意されたデータを用いて学習を行い、所望の性能が得られることを十分に検証した上で運用を開始することが多い。しかし、運用を続けていくと前述したような様々な要因によりデータ変化が生じ、性能が劣化する可能性がある。

MLOps基盤では、このような変化をいち早く検知するため、AIサービスへの入出力データの収集・蓄積及び、カスタマイズ可能なダッシュボードで可視化するモニタリング機能を提供する。また、しきい値を設定することでアラートメールを運用者や管理者に自動送信でき、監視に必要なリソースを削減するとともに迅速なデータ変化の検知を可能にした。

また、データの変化を検知しても、その変化の要因分析、及びAIモデルの更新や設備の設定変更などの対応判断には、高度なドメイン知識とAIモデルに対する深い理解が必要である。これらを自動化するAIモニタリング自体も盛んに研究されている分野であり⁽⁷⁾、東芝グループも、様々なAIモニタリングのアルゴリズムを開発している。これらをコンテナとして実装しておくことで、基盤上に取り込めるようにした。これにより最先端のモニタリング技術をタイムリーに適用し、監視・判断に掛かる人的リソースを大きく削減することが期待できる。

2.4 学習データ管理機能

AIモデルの性能や品質は、学習に使用したデータに依存する。AIモデルの更新を判断する際など、これらを検証するためには、学習データは、全てバックアップされているか、再現できることが望ましい。試行錯誤を含め、学習は繰り返し実施されるため、全データのバックアップには、リソース上

の課題がある。そこで、データのバージョン管理により過去の状態を再現する機能と、学習データとそのデータソースのメタデータを管理するデータカタログ機能を兼ね備えた、学習データ管理機能を開発した。この機能を利用することで、試行錯誤の履歴やデータ管理が自動化され、学習データのトレーサビリティと再現性が担保されるとともに、AIモデルの性能・品質の検証が確実かつ容易に実施可能となる。

2.5 CI/CD機能

MLOps基盤は、前述したパイプラインコンポーネントやAIモデルに対するCI/CD機能を提供する。CI/CD機能の導入によって、AIモデル学習と運用の定型処理が自動化され、迅速な開発運用が可能になる。具体的には、ソースコード管理ツールとの連携をサポートし、AIのアルゴリズムや各種コンポーネントのソースコード更新をトリガーとして、ビルドや、テスト、デプロイなどを自動実行できる。これにより、研究・開発の成果をスムーズに現場へと展開することが可能になる。

2.6 MLOps基盤の自動構築

前述した機能群を持ったMLOps基盤の構築を容易に実現するため、インフラを自動構築・管理するスクリプトとしてIaC (Infrastructure as Code)を開発した。IaCを用いることで、ユースケースに応じたITインフラの必要情報を設定するだけで、ネットワークの構築、インスタンスの立ち上げ、基盤アプリケーションのインストール、セキュリティ・権限の設定まで、全て自動で実行される。そのため、独自の環境に基盤を構築したいケースにおいても、難なく利用できる。

3. 適用事例

開発したMLOps基盤は、複数の事業に適用し、そこで発見された課題を基に更なる開発を続けている。ここでは、データタイプの異なる二つの適用事例について述べる。

3.1 時系列データによるプラント異常予兆検知

プラント内の複数のセンサーから得られる時系列データを用いて、異常検知AIアルゴリズムでプラント内設備の故障や異常の予兆を早期に発見する実サービスにおいて、季節変動や設備・センサーの修理交換によりデータが変化するためAIの精度を維持できないことが問題であった。この問題を解決するため、MLOps基盤の適用を行った。

基盤適用前は、AIが検知した異常予兆に対し、本当の異常予兆かデータ変化による誤検知かを判断するのは困難だったが、モニタリング機能でデータ変化を自動検知しヒストグラムや変化度合いを可視化することで、分析が容易になった。変化度合いの算出アルゴリズムとしても東芝グルー

独自のアルゴリズムを導入したことで、データの傾向変化をより適切に捉えることができた。また、パイプライン管理機能により学習させたいデータの期間を指定するだけでAIモデル作成と評価が可能となり、AIエンジニアでなくてもデータ変化に追従するAIモデルを作成し精度を維持できる。センサーデータを変更するだけでほかのプラント向けのAIモデルを作成することも可能であり、効率的な適用範囲の拡大も実現した。

3.2 画像データによる製品の自動検査

製品や部品の外観画像からAIが傷や色むらなどの問題を検出し、良品・不良品の判別を行う自動検査AIの運用において、最初に検査装置に導入されたAIモデルは目標としていた精度を達成していたが、撮影時の照明の明るさや微妙な色合いの変化によって精度が劣化する傾向が見られた。その度にAIエンジニアが画像を分析し、AIモデルの再学習・検証を行う必要があったため、運用に多大な人的リソースが掛かるという問題があった。

この問題を解決するためにMLOps基盤を導入し、AIの専門家でなくてもAIモデルの更新プロセスの実行を可能にした。また、モニタリング機能によって入力データの特徴量を可視化したことで問題の分析が容易になり、全体として運用工数を大きく削減できた。画像の取得先を変更するだけで、別の工場や製品についても同様の運用が可能であるため、今後は対象を拡大しつつ、それらを一つの基盤で一元管理することで、更なる運用及び研究・開発効率向上を目指している(図3)。

4. あとがき

IIoTサービスでのAI活用・運用をサポートするMLOps基盤を開発し、複数のIIoTサービスに適用した。これによ

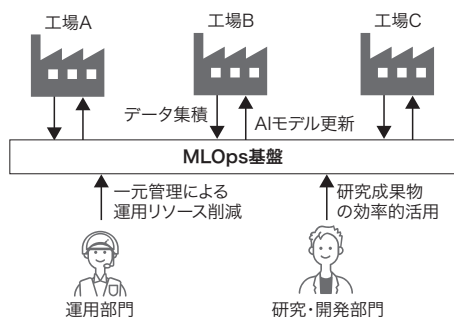


図3. MLOps基盤導入によるAIサービスの一元管理

MLOps基盤を導入することで、複数工場のAIサービスを一元管理して運用リソースを削減できるようになり、研究成果物も効率的に活用できる。

Centralized management of AI services by introduction of MLOps platform

り、AIの継続的な運用を効率化し、また東芝グループで開発したAI技術の試用・展開も容易になる。今後は、UI (User Interface) 統合による利便性向上やコスト低減など、現場適用で見えてきた課題を基に改善を続けるとともに、共通基盤として適用先を広げることで、スケールメリットによる更なる効率化を進めていく。

文献

- (1) Stanford Institute for Human-Centered Artificial Intelligence. Artificial Intelligence Index Report 2022. Stanford University, 2022, 229p. <https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf>, (accessed 2023-02-21).
- (2) Khalil, R.A. et al. Deep Learning in the Industrial Internet of Things: Potentials, Challenges, and Emerging Applications, IEEE Internet Things J. 2021, 8, 14, p.11016-11040.
- (3) 東芝のAI <<https://www.global.toshiba/jp/cps/corporate/ai.html>>, (参照 2023-02-21).
- (4) WIPO Technology Trends 2019: Artificial Intelligence, World Intellectual Property Org., Geneva, Switzerland, 2019, p.60, <https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf>, (accessed 2023-02-21).
- (5) Kreuzberger, D. et al. Machine Learning Operations (MLOps): Overview, Definition, and Architecture. arXiv 2022, <<https://arxiv.org/ftp/arxiv/papers/2205/2205.02302.pdf>>, (accessed 2023-02-21).
- (6) Klaise, J. et al. Monitoring and explainability of models in production. 2020, arXiv:2007.06299. <<https://arxiv.org/pdf/2007.06299.pdf>>, (accessed 2023-02-21).
- (7) Hendrycks, D. et al. A BASELINE FOR DETECTING MISCLASSIFIED AND OUT-OF-DISTRIBUTION EXAMPLES IN NEURAL NETWORKS. 2018, arXiv:1610.02136. <<https://arxiv.org/pdf/1610.02136.pdf>>, (accessed 2023-02-21).



小島 知也 KOBATAKE Tomoya
デジタルイノベーションテクノロジーセンター
技術開発室 サービスプラットフォーム開発部
Service Platform Development Dept.



山田 正隆 YAMADA Masataka
研究開発センター
知能化システム技術センター AI基盤技術開発部
AI Platform Technology Dept.