

## 東芝アナリティクスAI “SATLYS”の ディープラーニングモデル開発プラットフォーム

Deep Learning Model Development Platform for SATLYS Toshiba Analytics AI Services

西澤 実 NISHIZAWA Minoru 後藤 哲也 GOTOU Tetsuya 江藤 雅哉 ETO Masaya

産業分野では、IoT (Internet of Things)、ビッグデータ処理、及びAIに関する技術開発が活発化している。しかし、標準的なAIモデルの開発手法が確立していない、AI技術者の能力や経験が異なる、開発ツールが多様化している、などの問題も顕在化してきている。

このような環境の中、東芝デジタルソリューションズ(株)は、東芝IoTアーキテクチャー“SPINEX”により、様々な顧客向けにデジタルトランスフォーメーションの実現に取り組んでいる。そして、SPINEXを支える新たなプロフェッショナルサービスとして、2017年10月にAI技術を活用して解析を行う東芝アナリティクスAI“SATLYS”の提供を開始した。SATLYSでは、高品質なアナリティクスAIサービスを提供するため、ディープラーニングモデルの開発で得た知見や独自技術を蓄積・共有するためのプラットフォームを構築した。これにより、開発環境の整備に要するコストの削減や、独自技術の適用によるモデルの精度向上などが可能になり、顧客のニーズに応じた高品質で高効率なサービスを開発できる。

Efforts to develop a wide variety of technologies related to the Internet of Things (IoT), big data, and artificial intelligence (AI) have been actively promoted in the industrial field in recent years. However, this trend is now encountering various issues including the lack of an established standard methodology for the development of AI models, human resources having different experience levels and skills, and diverse development tools.

Under these circumstances, Toshiba Digital Solutions Corporation is engaged in activities to realize digital transformation that is of high value to individual customers by means of its SPINEX Toshiba IoT architecture. As part of these activities, we released SATLYS, a suite of AI analytics services, in October 2017 as a set of new professional services to perform analyses using AI technologies. In order to offer high-quality AI analytics services, we have constructed a platform to store and share the knowledge and proprietary technologies that we have accumulated through the development of deep learning models. This platform makes it possible to reduce the costs of constructing development environments and improve the accuracy of models, resulting in the creation of services with high quality and high efficiency in response to customers' requirements.

### 1. まえがき

近年、産業分野では、IoT (Internet of Things) 技術でビッグデータを収集して可視化し、更にAI技術を活用することで、熟練者の知見をデジタル化して生産性や業務効率を改善する取り組みが活発化している。

東芝デジタルソリューションズ(株)は、東芝IoTアーキテクチャー“SPINEX”を支える新たなAIサービスとして、2017年10月に東芝アナリティクスAI“SATLYS”<sup>(1)</sup>の提供を開始した。SATLYSは、顧客との共創を通して、既製のAIモデル自動生成サービスなどでは賄えないような、高度なAIモデルの設計、学習、及びAI推論サービスの構築を行い、検査データや、センサーデータ、業務データ、行動データなどの解析により、生産性や業務効率の改善を提供するプロフェッショナルAIサービスである。ディープラーニ

ング技術に代表される最先端のAI技術を駆使してIoTデータを解析することで、高精度な識別や、予測、要因推定、異常検知、故障予兆検知、行動推定などを行う<sup>(2)</sup>。

ここでは、ディープラーニング技術を用いて開発したモデルの知見を蓄積し、高品質なアナリティクスAIサービスを効率良く開発するための、ディープラーニングモデル開発プラットフォームについて述べる。

### 2. SATLYSの開発プロセス

#### 2.1 開発プロセスの概要

SATLYSの開発プロセスの概要を図1に示す。

このプロセスでは、まず、デジタルコンサルテーションを通して、顧客の業務課題の抽出、KPI (Key Performance Indicator) の設定、AI導入の提案を行う。そして、これらを基に、サンプルデータを用いたモデル設計、大規模デー

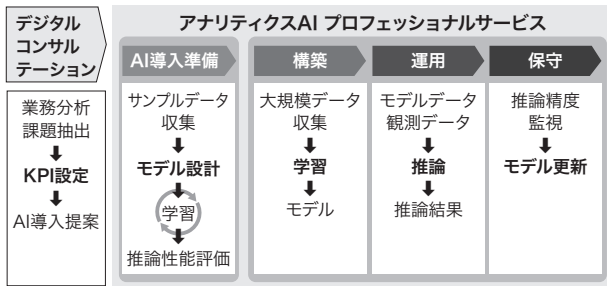


図1. SATLYS開発プロセスの流れ

アナリティクスAIサービスの開発プロセスは、モデル設計、モデルの学習、AI推論サービスの構築、運用、保守などから成り、これらをプロフェッショナルサービスとして提供する。

Processes for development using SATLYS

タを用いたモデルの学習、オンプレミス環境やクラウド環境でのAI推論サービスの構築、運用、更には推論モデルの精度監視、モデル更新などを、プロフェッショナルサービスとして提供する。

## 2.2 アナリティクスAIサービス開発における問題

アナリティクスAIサービス開発では、以下の問題が挙げられる。

- (1) 標準的なモデル開発手法が未確立 産業分野におけるアナリティクスAIサービスは、今後、生産性や業務効率の改善の中核として重要である。そして、優れたAI技術による魅力あるサービスをいかに早く現場に提供できるかが、企業の競争力や活動成果を左右する。しかし、これまでは、標準的なモデルの開発手法が確立されておらず、試行錯誤を伴っていた。短納期でのサービス開発を実現するには、可能な限り試行錯誤を減らす施策が必要である。
- (2) AI技術者の能力や経験の違い アナリティクスAIサービス開発では、システム開発を行う従来の技術者と、高度な専門性を持つAI技術者との連携が必要となる。一方で、AI技術者は不足しており、技術者の育成やパートナーの活用などが行われている。この状況で、能力や経験の異なるAI技術者が開発を行うことから、設計や成果物の品質のばらつきが生じやすくなる。したがって、AI技術者のスキルによらず開発を行い、一定の品質でサービスを提供できること、また技術やノウハウの共有によりAI技術者をレベルアップする仕組みが必要である。
- (3) モデル開発手法の多様化への対応と新技術の取り込み 近年、モデル開発などで利用するOSS（Open

Source Software）の統計解析や機械学習などの数理統計ライブラリー、ディープラーニングのライブラリー、データ可視化支援ツールなどが急速に増加している。科学技術分野では、ディープラーニングを用いた新たな開発アプローチや方法論が日進月歩で進化しているので、最新のツールや技術をいち早く導入するとともに、常にその時点で最適な手法を選択してビジネスに適用することが必要である。

## 3. ディープラーニングモデル開発プラットフォームの概要

2章で述べたように、アナリティクスAIサービス開発には、モデル開発方法の未確立、開発メンバーのスキルの違い、モデル開発手法の多様化などの問題があり、これらについては、以下の解決策が有効である。

- (1) モデル学習・推論のソフトウェア構造の標準化
- (2) モデル開発ノウハウの再利用
- (3) モデル開発環境の標準化
- (4) モデル開発手法の標準化

### 3.1 構成要素

前述の解決策を実現する基盤技術として、ディープラーニングモデル開発プラットフォームを構築した。ディープラーニングモデル開発プラットフォームは、図2に示すように、四つの構成要素から成る。

- (1) モデル開発フレームワーク
- (2) モデル開発コンポーネント
- (3) モデル開発環境
- (4) モデル開発標準

これらの個々の構成要素について、以下に説明する。

### 3.2 モデル開発フレームワーク

モデル開発フレームワークは、ディープラーニングモデルの学習と推論の各処理を実装するモジュールのインターフェースを定義する。ここで、ディープラーニングモデルとは、多層構造のニューラルネットワークである。ニューラルネットワークは、入力層、出力層、隠れ層から構成され、各層は複数個のノードを持つ。層と層の間は、ノード同士のつながりの強さを表す重み付きのエッジで結ばれている。ディープラーニングモデルの学習と推論の処理は、図3の(1)から(5)に示すような、典型的な処理手順がある。

- (1) クレンジング 異質なデータを取り除き、誤りの補正や不足データの補間など、データの整理を行う。
- (2) 前処理 入力データの値が0から1の範囲に収ま

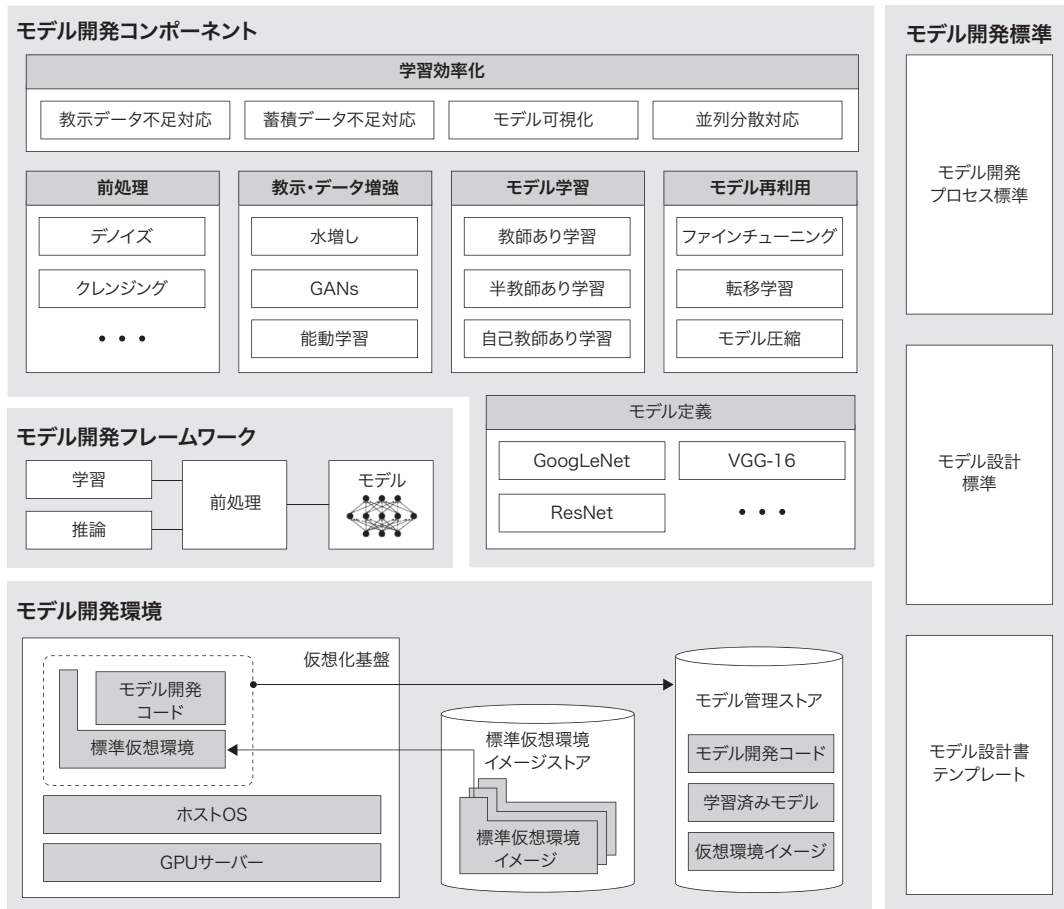


図2. ディープラーニングモデル開発プラットフォームの構成

モデル開発フレームワーク、モデル開発コンポーネント、モデル開発環境、モデル開発標準の四つの要素で構成される。

Configuration of deep learning model development platform

るように、その値を正規化する処理などを行う。

- (3) 推論 入力データの種類(画像データ、センサーデータなど)や分析問題(回帰、分類など)に合わせて定義したディープラーニングモデルに対し、入力データを与え、推論結果を出力する。
- (4) モデル評価 モデルからの出力データと、入力データに対する教師データ(正解)との誤差を計算する。
- (5) 重み更新 計算した誤差が小さくなるように、各エッジの重みを少しずつ適正な値に更新することを繰り返し、最適なモデルを求める。

このような典型的な処理を、開発標準の枠組みの形で提供することにより、設計・開発プロセスを均一化する。学習と推論の両方で共通に使用するモデル定義や前処理のモジュールは、両処理から分離し共通化することで、保守性や拡張性を向上させる。

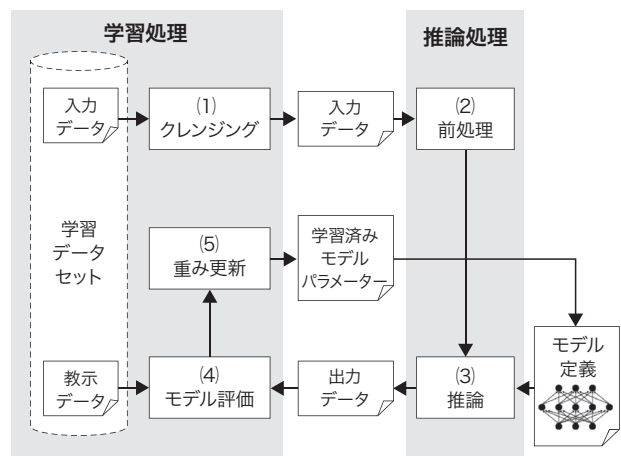


図3. ディープラーニングモデルの学習・推論処理フロー

学習処理は、クレンジングから始まり、前処理、推論、モデル評価、重み更新を繰り返して最適な処理を求めるフローである。

Flow of processes for training and inference in deep learning model

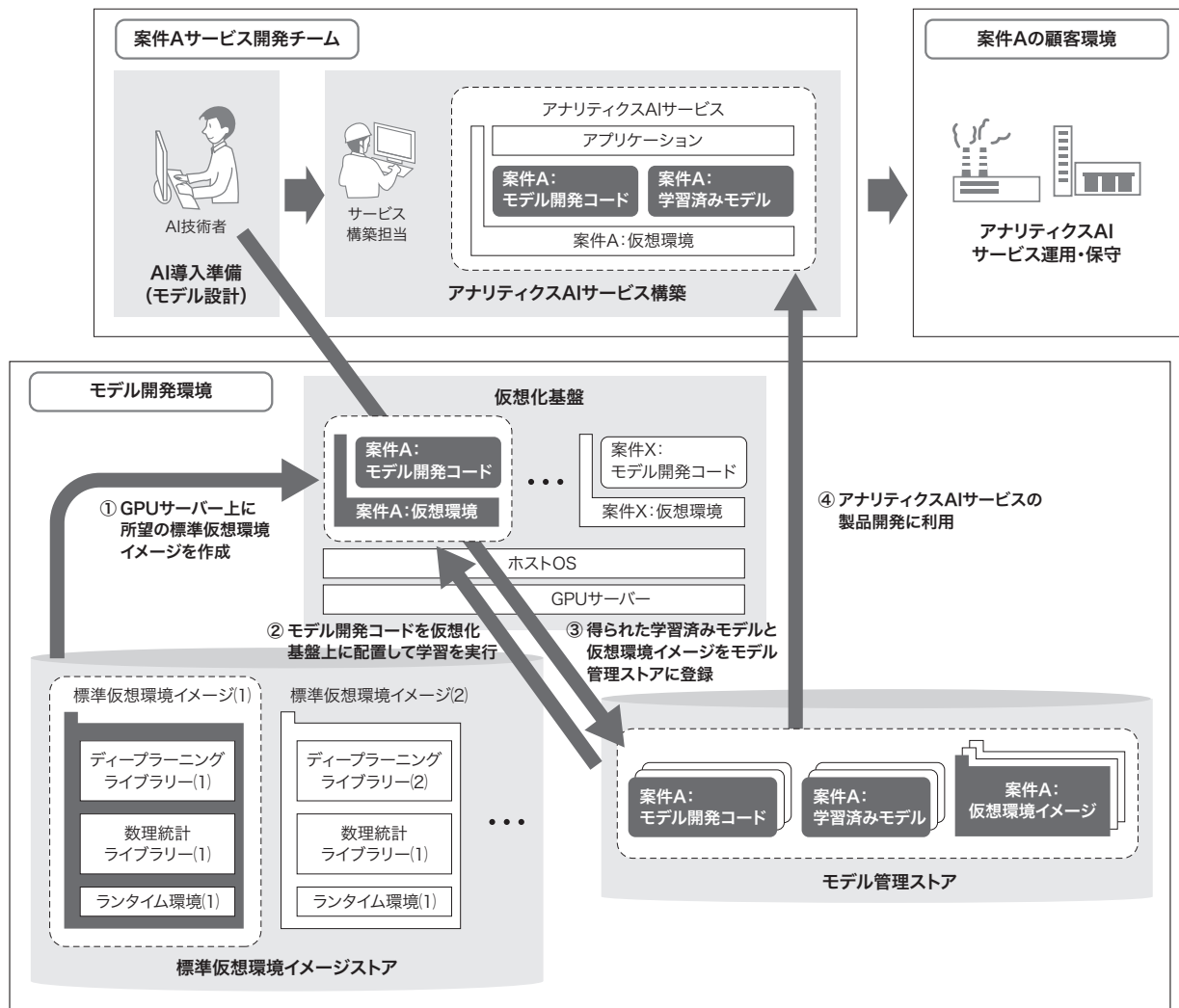


図4. モデル開発環境を使ったモデル設計の流れ

モデル開発コードを仮想化基盤上に配置し、モデル学習の実行から学習済みモデルのモデル管理ストアへの登録に至る一連の処理が全自動で実行でき、ディープラーニング実行時の手間を減らすことができる。

Flow of model design using model development environments

### 3.3 モデル開発コンポーネント

モデル開発コンポーネントは、前処理、教示・データ増強、ディープラーニング技術によるモデル学習、学習済みモデルの再利用など、これまでのモデル開発で実績のある各学習ステージのコア技術をアセット化したソフトウェア部品群である。著名なニューラルネットワーク構造を実装したモデル定義(“GoogLeNet”, “VGG-16”, “ResNet”など)や、GANs (Generative Adversarial Networks: 敵対的生成ネットワーク)をベースに、生成画像と実画像の違いを無くすための特徴抽出器と特徴識別器を追加した独自の画像生成技術<sup>3)</sup>が一例である。加えて、取り組む対象の問題特性を考

慮して、適切なコンポーネントを選択でき、それらを組み合わせることにより、効率的な学習を実現できる。このようにディープラーニングによるモデル開発の知見や独自技術をソフトウェア部品として蓄積し、活用していくことで、新規開発部品を削減することによる効率化や、AI技術者のスキルに依存しない品質の確保、そして早期のサービス提供を可能にする。

### 3.4 モデル開発環境

モデル開発環境は、ディープラーニングを使ってモデル開発を行うための仮想化環境である。ディープラーニングを実行するには、高速な学習処理を可能とするGPU (Graph-

ics Processing Unit) を搭載したGPUサーバーを用意し、そのOS (基本ソフトウェア) 上に、GPUドライバーや、ラuntime環境、統計・機械学習ライブラリー、ディープラーニングのライブラリーなどをインストールした環境を用意する必要がある。ディープラーニングのOSSライブラリーは、種類が多く、採用するライブラリーやそのバージョンに応じて、必要なソフトウェア構成も異なる。

そこで、ディープラーニングモデル開発プラットフォームでは、仮想化基盤を利用して、モデル開発環境の標準化を行っている。モデル開発環境の標準化では、開発実績のあるソフトウェア構成を標準の仮想環境イメージとして事前に用意することで、サービス開発の度に発生する環境構築作業を省略し、実績のある環境での開発を可能にする。

また、このモデル開発環境では、図4に示すように、ディープラーニング実行時の手間を減らすために、①GPUサーバー上に所望の標準仮想環境イメージを作成、②モデル管理ストアに保存されたモデル開発コードを仮想化基盤上に配置して学習を実行、③得られた学習済みモデルと仮想環境イメージをモデル管理ストアに登録する、の一連の作業を全自動で実行できるようにした。これにより、複雑な実行作業が簡略化され、人的なミスを削減できる。

### 3.5 モデル開発標準

モデル開発標準は、以下の要素から構成される。

- (1) モデル開発プロセス標準 モデルの開発手順と、各手順で作成されるべき成果物の定義である。開発手順の各作業項目は、作業項目名、目的、開始条件、終了条件、入力となる成果物、作業内容、出力となる成果物、利用するモデル設計書テンプレートにより定義される。
- (2) モデル設計書テンプレート モデル設計書の目次構成や、記述すべき内容を定義したドキュメントテンプレートから成る。開発プロセスの作業ごとに作成すべき成果物が決まっており、それらのテンプレートが定義されている。
- (3) モデル設計標準 前記(1)のモデル開発プロセス標準の作業項目とその成果物について、事例に基づいて具体的に説明したガイドラインである。

これら三つの構成要素により、モデル開発における試行錯誤の削減、及び設計作業における抜けや漏れの回避ができる。また、ディープラーニングモデル開発プラットフォームのほかの構成要素で、利用すべきポイントを定義すれば、モデル開発ノウハウの蓄積とディープラーニングモデル開発

プラットフォームの利用促進が図れる。

## 4. あとがき

アナリティクスAIサービス開発において、モデル開発手法の未確立、AI技術者の能力や経験などのスキルの違い、モデル開発手法の多様化といった問題を解決するための、ディープラーニングモデル開発プラットフォームについて述べた。今後は、学習効率化に向けてモデル開発コンポーネントを拡充し、更なるサービス開発の品質及び生産性の向上を図っていく。

## 文献

- (1) 東芝デジタルソリューションズ, “AIによるビジネス変革を加速する東芝アナリティクスAI「SATLYS™」を提供開始”. ニュースリリース. <<https://www.toshiba-sol.co.jp/news/detail/20171030.htm>>, (参照 2018-07-25).
- (2) 東芝デジタルソリューションズ, “SATLYSを支える、分析メソッドと分析技術”. DiGiTAL T-SOUL. 2018, 24, #03. <<https://www.toshiba-sol.co.jp/articles/tsoul/24/003.htm>>, (参照 2018-07-25).
- (3) 伊藤秀将, ほか. 画像認識精度を向上させるディープラーニング技術を用いた学習用画像の自動生成. 東芝レビュー. 2017, 72, 4, p.18-21. <[https://www.toshiba.co.jp/tech/review/2017/04/72\\_04pdf/a05.pdf](https://www.toshiba.co.jp/tech/review/2017/04/72_04pdf/a05.pdf)>, (参照 2018-07-25).



西澤 実 NISHIZAWA Minoru  
東芝デジタルソリューションズ (株)  
ソフトウェア& AI テクノロジーセンター  
ディープラーニング技術開発部  
Toshiba Digital Solutions Corp.



後藤 哲也 GOTOU Tetsuya  
東芝デジタルソリューションズ (株)  
ソフトウェア& AI テクノロジーセンター  
システム&サービス技術部  
Toshiba Digital Solutions Corp.



江藤 雅哉 ETO Masaya  
東芝デジタルソリューションズ (株)  
ソフトウェア& AI テクノロジーセンター  
ディープラーニング技術開発部  
Toshiba Digital Solutions Corp.