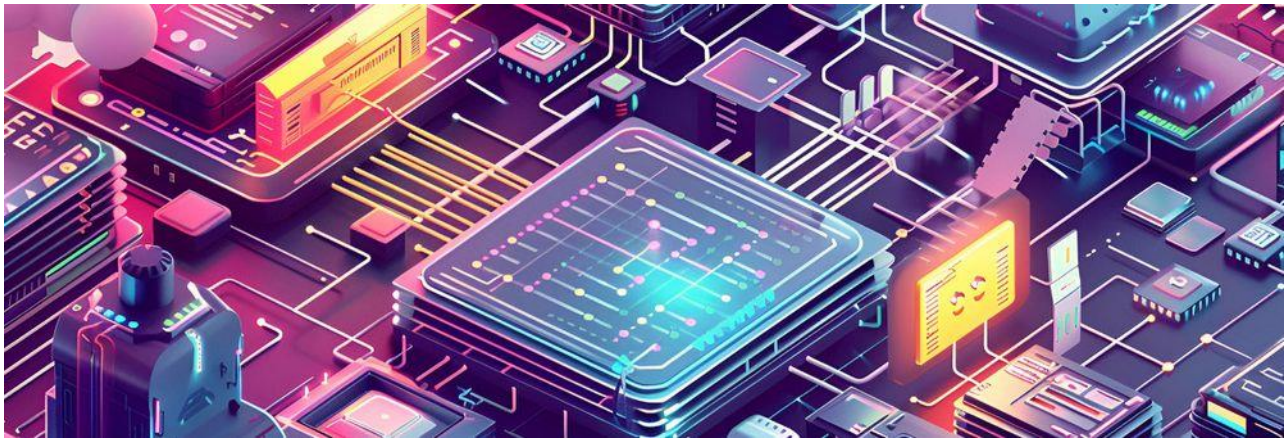
インテル® Tiber™ AI クラウド上での ML ワークロードの構築と開発

この記事は、インテルのブログで公開されている「Build and Develop ML workloads on Intel® Tiber™ Developer Cloud」の日本語参考訳です。原文は更新される可能性があります。原文と翻訳文の内容が異なる場合は原文を優先してください。



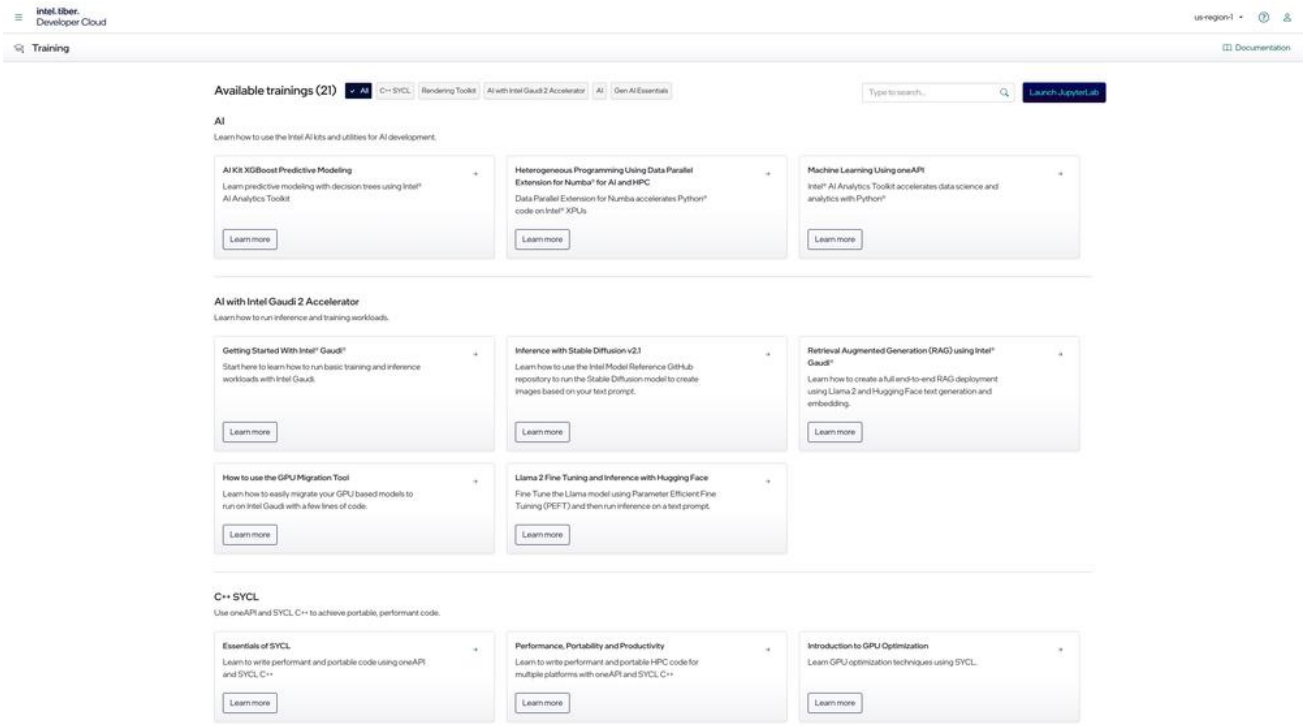
マシンラーニング (ML) (英語) は、人工知能 (AI) の一分野であり、データから学習し、明示的な指示なしに未知のデータに対して新しい結果を予測する統計的アルゴリズムを使用します。ML モデルのトレーニングには、大量のデータ、計算能力、インフラストラクチャーが必要です。インテルは、開発者が最新のインテル® ハードウェア上でインテルの最適化されたソフトウェアを使用して AI 開発を高速化できる柔軟なソリューション、インテル® Tiber™ AI クラウド (旧称: インテル® Tiber™ デベロッパー・クラウド) を提供しています。

インテルはまた、oneAPI ライブラリーを使用して主要な ML フレームワークを最適化して、インテル® アーキテクチャー全体で最高のパフォーマンスを発揮できるようにしています。これらのソフトウェア最適化は、同じフレームワークの標準実装よりも優れたパフォーマンスを実現するのに役立ちます。インテル® Tiber™ AI クラウドは、ML フレームワークを利用する AI アプリケーションとソリューションに、インテル® Gaudi® 2 AI アクセラレーターやインテル® Xeon® スケーラブル・プロセッサーを含む幅広いハードウェアへのアクセスを提供します。開発者は任意の CPU や GPU 上でワークロードを学習、プロトタイプ化、テスト、実行することができ、無料の Jupyter* Notebook とチュートリアル (英語) を通じてプラットフォームとソフトウェアの最適化をテストできます。

この記事では、インテル® Tiber™ AI クラウド上で ML ワークロードを実装および開発するベスト・プラクティスを紹介します。この記事の手順を実行する前に、インテル® Tiber™ AI クラウドを使い始めるガイドをお読みになることを推奨します。

インテル® Tiber™ AI クラウドの一般的な使用手順:

1. cloud.intel.com (英語) に
2. サインインします。まだアカウントをお持ちでない場合は、こちらの手順に従ってアカウント登録を行ってください。
3. サインインしたら、コンソールの左上のメニューアイコンをクリックして Training (英語) を選択します。
4. 受講するトレーニングの **[Launch]** ボタンをクリックします。



JupyterLab* では、開発者のニーズに応じていくつかのカーネルタイプが用意されています。カーネルは事前にインストールされた Python* 環境であり、ユーザーが新しいノートブックを開くと、JupyterHub* は指定された環境に対応するパッケージのインストールを検出します。ほとんどの場合、以下のサンプルコードの実行に必要なパッケージは、ベースカーネルに含まれています。

インテル® Tiber™ AI クラウドでマシンラーニングを始める

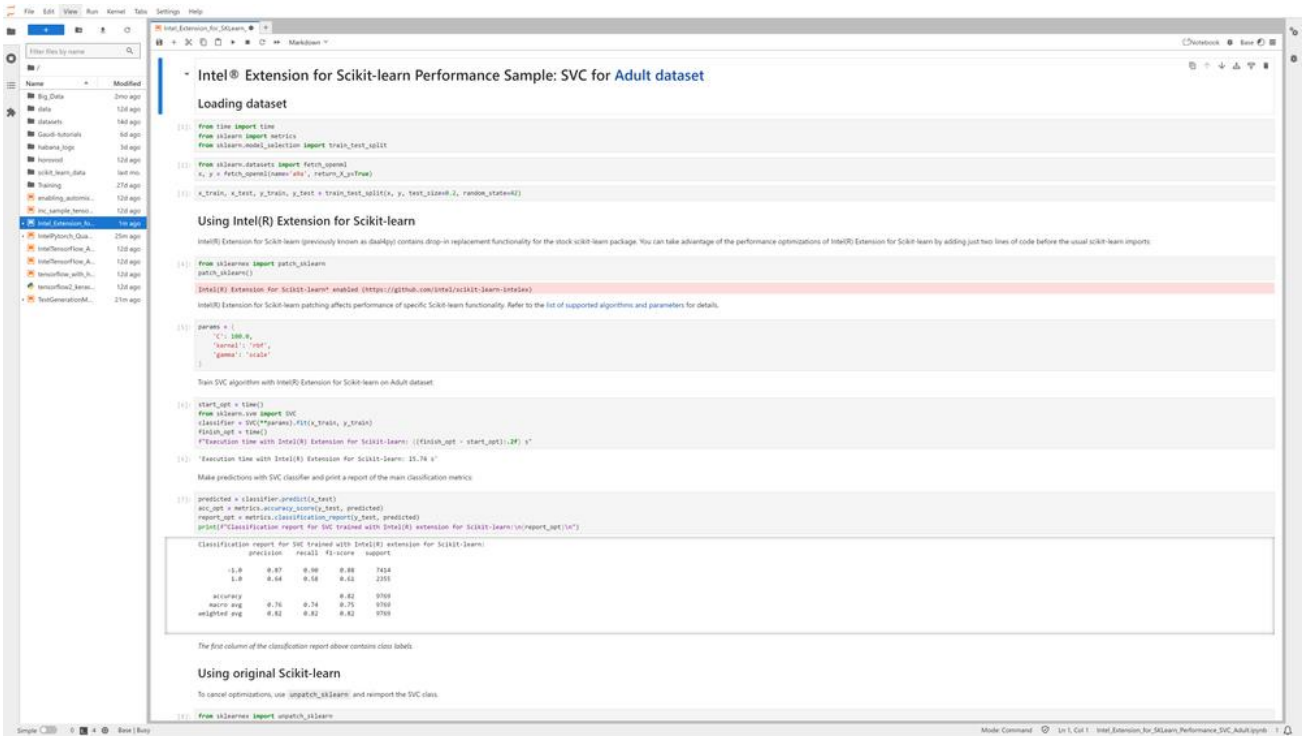
scikit-learn* 向けインテル® エクステンション: scikit-learn* (sklearn) は、予測データ分析およびマシンラーニングに有用なシンプルで効率的な Python* パッケージです。scikit-learn* 向けインテル® エクステンション (英語) は、インテル® アーキテクチャー上で scikit-learn* の多くのアルゴリズムのパフォーマンスを向上します。シングルノードとマルチノードの両方で、インテル® アーキテクチャー上の ML アルゴリズムのパフォーマンスを最大限に引き出します。

以下は、インテル® Tiber™ AI クラウドで scikit-learn* 向けインテル® エクステンションの成人データセット向け SVC サンプル (英語) を実行する手順です。

1. JupyterLab* を起動します。
2. ダッシュボードのメニューから **[File] > [Open from URL...]** を選択し、以下の URL をコピーしてペーストします。
   https://raw.githubusercontent.com/oneapi-src/oneAPI-samples/master/AI-and-Analytics/Features-and-Functionality/Intel_Extension_For_SKLearn_Performance_SVC_Adult/Intel_Extension_for_SKLearn_Performance_SVC_Adult.ipynb

3. **[Kernel] > [Change Kernel] > [Base]** を選択して、Base カーネルに変更します。
4. サンプルコードのすべてのセルを実行し、出力を確認します。

このサンプルコードは、scikit-learn* 向けインテル® エクステンションを利用して SVC アルゴリズムでトレーニングと予測を行います。また、scikit-learn* 向けインテル® エクステンションと scikit-learn* のパフォーマンスを比較し、scikit-learn* のパッチが元の scikit-learn* よりも大幅なパフォーマンス向上をもたらすことを確認できます。



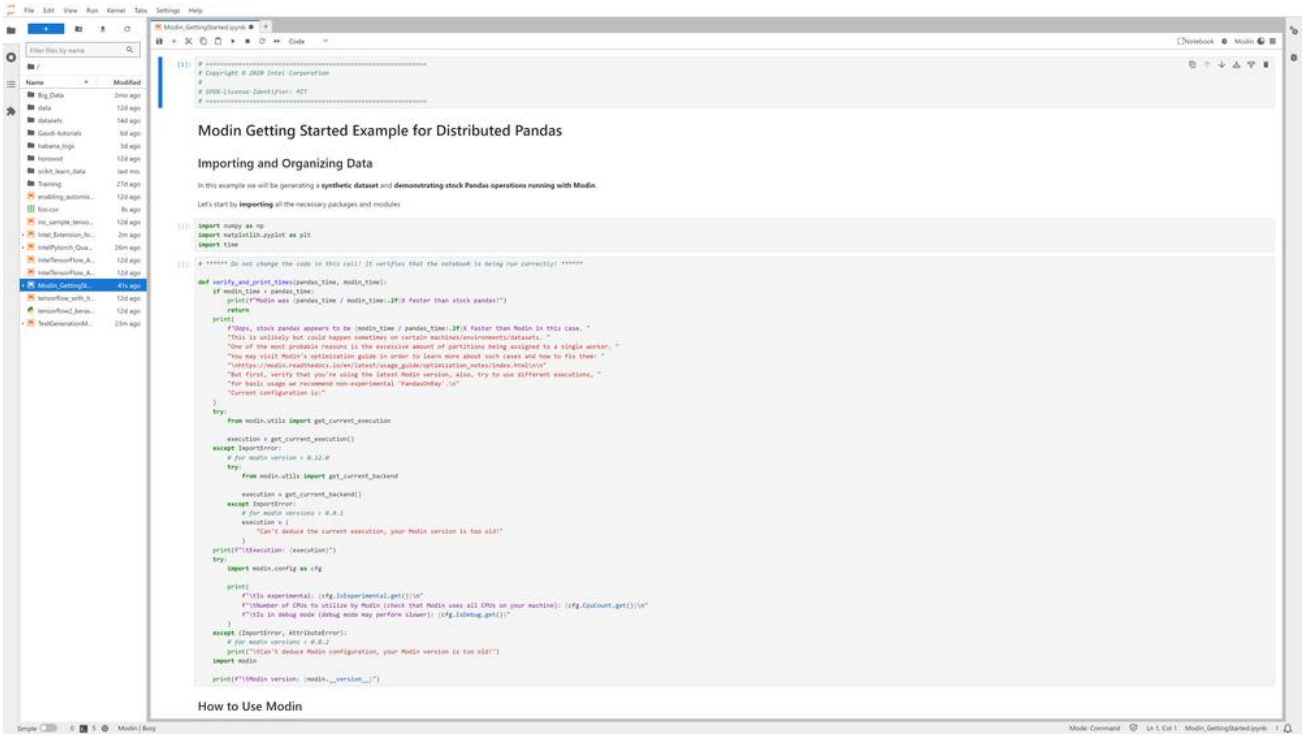
Modin: Modin (英語) は pandas の代替です。Modin を使用することで、データサイエンティストは API コードを変更することなく、データ分析に集中できます。このディストリビューションは、インテル® ハードウェア上で処理を加速する最適化を追加します。インテルは、すべての最適化をオープンソースの Modin にアップストリームしています。

以下は、インテル® Tiber™ AI クラウドで Modin の入門サンプル (英語) を実行する手順です。

1. JupyterLab* を起動します。
2. ダッシュボードのメニューから **[File] > [Open from URL...]** を選択し、以下の URL をコピーしてペーストします。
   https://raw.githubusercontent.com/oneapi-src/oneAPI-samples/master/AI-and-Analytics/Getting-Started-Samples/Modin_GettingStarted/Modin_GettingStarted.ipynb

3. **[Kernel] > [Change Kernel] > [Modin]** を選択して、Modin カーネルに変更します。
4. サンプルコードのすべてのセルを実行し、出力を確認します。

このサンプルコードは、Modin により加速された pandas 関数を使用し、Modin と標準 pandas 関数のパフォーマンスを比較します。



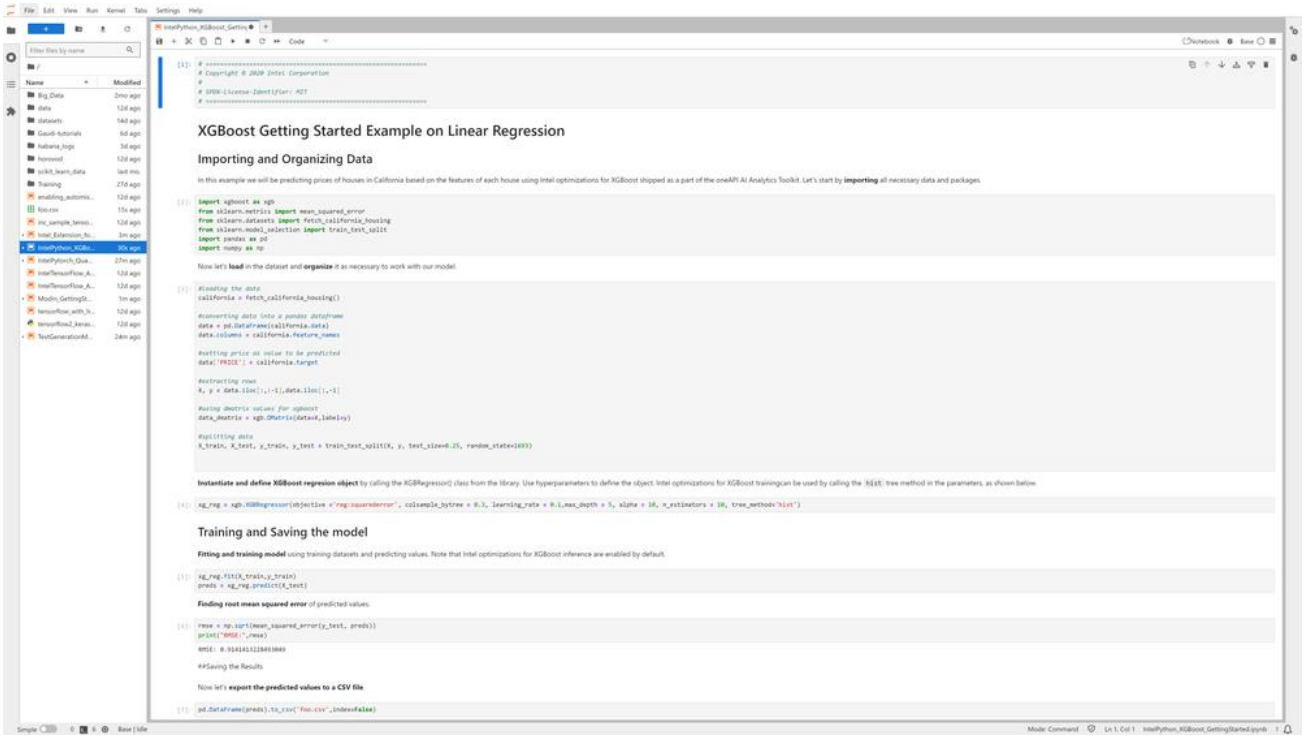
インテルによる勾配ブースティング最適化: 勾配ブースティングは、複数のモデルからの予測を組み合わせて堅牢な予測モデルを構築する ML アンサンブル手法です。XGBoost (Extreme Gradient Boosting) (英語) は、スケーラブルな分散勾配ブースティング決定木を実装するオープンソースの ML ライブラリーです。ユーザーは、インテルが開発した daal4py ライブラリーの高速なツリー推論機能を使用することで、精度を損なうことなく勾配ブースティング推論を加速できます。

以下は、インテル® Tiber™ AI クラウドで XGBoost の入門サンプル (英語) を実行する手順です。

1. JupyterLab* を起動します。
2. ダッシュボードのメニューから **[File] > [Open from URL...]** を選択し、以下の URL をコピーしてペーストします。
   https://raw.githubusercontent.com/oneapi-src/oneAPI-samples/master/AI-and-Analytics/Getting-Started-Samples/IntelPython_XGBoost_GettingStarted/IntelPython_XGBoost_GettingStarted.ipynb

3. **[Kernel] > [Change Kernel] > [Base]** を選択して、Base カーネルに変更します。
4. サンプルコードのすべてのセルを実行し、出力を確認します。

このサンプルコードは、インテルによる XGBoost 最適化を使用して XGBoost モデルを設定し、トレーニングし、データセットの特徴に基づいて予測します。



インテルにより最適化された上記のフレームワークはすべて、AI ツールの一部として入手できます。

インテル® Tiber™ AI クラウドで最新のシリコン・ハードウェアと最適化されたソフトウェアにアクセスし、革新的な AI プロジェクトを開発および強化してみてください。インテルの AI ツールとフレームワークの最適化 (英語) をチェックし、インテルの AI ソフトウェア・ポートフォリオ (英語) の基盤となる統一されたオープンな標準ベースの oneAPI プログラミング・モデルについて学びましょう。また、業界をリードする独立系ソフトウェア・ベンダー (ISV)、システム・インテグレーター (SIer)、OEM、エンタープライズ・ユーザーとのコラボレーション (英語) が、AI の導入をどのように加速させるかをご覧ください。

エクセルソフトでは、インテル® Tiber™ AI クラウドの評価用クーポンを配布しています。ご興味のある方は、こちらのフォームよりお問い合わせください。

関連資料

- インテルの AI 開発ツールとリソース (英語)
- oneAPI 統一プログラミング・モデル
- インテル® Tiber™ AI クラウド
- インテルの AI ツール
- 記事 - scikit-learn* 入門: ライブラリーとインテル® エクステンションの包括的なガイド (英語)
- 記事: Modin 入門: pandas を高速化する手順 (英語)
- 記事 - XGBoost 入門: ライブラリーとインテル® オプティマイゼーションの包括的なガイド (英語)
- マシンラーニング (英語)