ディープラーニング向け インテル® ソフトウェア製品の紹介 (I)

AI
人工知能

高度なプログラミングをすることなく、人間の心理と関連した認識機能を実行するために、経験を基にした機械の学習する能力

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

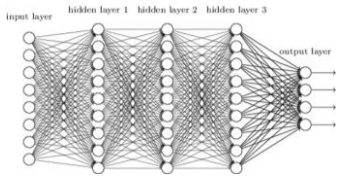
マシンラーニング技術の分析

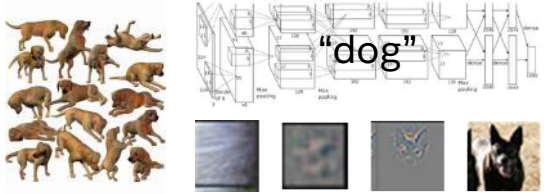
マシンラーニング
経験 (データ) から自主的に学ぶ計算手法

訓練
モデル構築のための訓練
- モデル構築時間は重要

推論
モデルを分類、予測、認識に適用
- 簡単に分散処理
- 基準: スループット、TCO @ scale

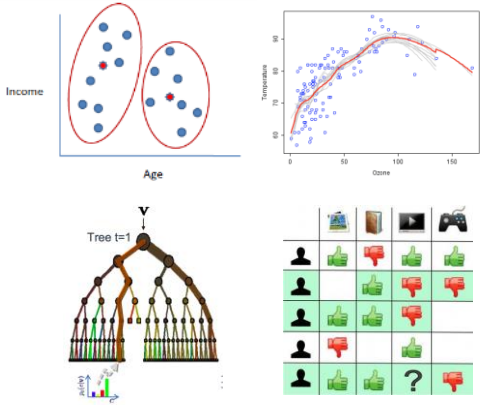
ディープラーニング
多くの隠れたレイヤーへの階層的アプローチ
– 画像や音声、自然言語などの正確に分類されたデータからフレームを得る。特徴は学習される。





代表的なユーザー: CSP、HPC

その他 (従来) の ML
極めて少ない (1 つまたは 2 つの) 隠れたレイヤーを使ってクラスタリング、回帰、分類を行う従来のマシンラーニングの手法。特徴エンジニアリングが必要。



代表的なユーザー: Enterprise、HPC

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

AI 人知工能

ソリューション
ソリューション・アーキテクト

人工知能 (ARTIFICIAL INTELLIGENCE)

AI ソリューション・カタログ (Public (英語) & Internal (英語))

プラットフォーム | 金融 | ヘルスケア | エネルギー | 産業 | 輸送 | 小売 | 家庭 | その他…

ツールキット
アプリケーション開発者

ディープラーニングのデプロイメント

OpenVINO™ ツールキット[1]

TensorFlow*、Caffe、MXNet* を使った CPU、プロセッサー・グラフィックス、FPGA & VPU 向けの推論デプロイメント用 Open Visual Inference & Neural Network Optimization ツールキット

インテル® Movidius™ SDK (英語)

TensorFlow* & Caffe を使ってすべてのインテル® Movidius™ VPU に最適化された推論エンジン

ディープラーニング

近日公開!

インテル® Deep Learning Studio

ディープラーニング開発サイクルを効率化するオープンソースのツール

ライブラリー
データ・サイエンティスト

マシンラーニング・ライブラリー

Python*
- scikit-learn (英語)
- pandas (英語)
- NumPy* (英語)

R
- Cart (英語)
- Random Forest (英語)
- e1071 (英語)

Distributed
- MLLib (Spark*) (英語)
- Mahout (英語)

ディープラーニング・フレームワーク

近日公開!

CPU に最適化済み

TensorFlow*、MXNet*、Caffe、BigDL/Spark* (英語)

最適化作業中

Caffe2 | PyTorch* | PaddlePaddle*

ファウンデーション
ライブラリー開発者

分析、マシンラーニング & ディープラーニング用プリミティブ

Python*[2]
マシンラーニングに最適化されたインテルのディストリビューション

インテル® DAAL
インテル® データ・アナリティクス・アクセラレーション・ライブラリー (マシンラーニング用)

インテル® MKL-DNN
clDNN
CPU とインテル® グラフィックスをサポートするオープンソースのディープ・ニューラル・ネットワーク関数

ディープラーニング・グラフ・コンパイラー

インテル® nGraph コンパイラー (Alpha)

複数のフレームワーク (TensorFlow*、MXNet*、ONNX*) を使用して複数のデバイス (CPU、GPU、NNP) に最適化されたディープラーニングのモデルを計算するオープンソースのコンパイラー

ハードウェア
IT システム・アーキテクト

AI ファウンデーション

ATOM | CORE 8th Gen | XEON PLATINUM

データセンター
端末
デバイス

ディープラーニング・アクセラレーター

2019 公開

NERVANA | STRATIX 10 | ARRIA 10 | MOVIDIUS | GNA

NNP L-1000 ← 推論 →

[1] 旧インテル® コンピューター・ビジョン SDK
[2] インテル® Distribution for Python*
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。
すべての製品、コンピューター・システム、日付、図は現在の予想に基づくものであり、予告なく変更されることがあります。

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

Ai.intel.com (英語)

インテル® ディープラーニング・ブースト

将来のインテル® Xeon® スケーラブル・プロセッサー (開発コード名 Cascade Lake)

インテル® Xeon® スケーラブル・プロセッサーでは数値精度の低い命令を使って組み込まれたディープラーニングのアクセラレーションが可能

推論

Q3' 2017-Q3' 2018

FP32 精度

Q3' 2018-Q4' 2018

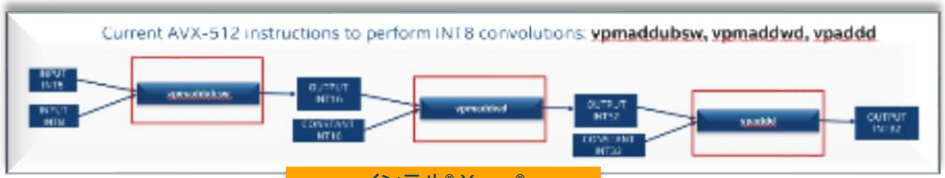
INT8 精度

インテル® Xeon® スケーラブル・プロセッサー (開発コード名 Skylake)

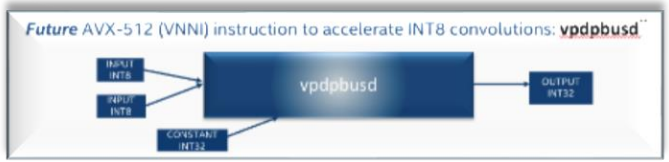
Q4' 2018- 2019

VNNI サポート INT8

将来のインテル® Xeon® スケーラブル・プロセッサー (開発コード名 Cascade Lake)



インテル® Xeon® スケーラブル・プロセッサー



将来のインテル® ディープラーニング・ブースト対応 インテル® Xeon® スケーラブル・プロセッサー

https://www.intel.com/content/www/us/en/now/your-data-on-intel/deep-learning-boost-video.html (英語)
https://software.intel.com/en-us/articles/lower-numerical-precision-deep-learning-inference-and-training (英語)

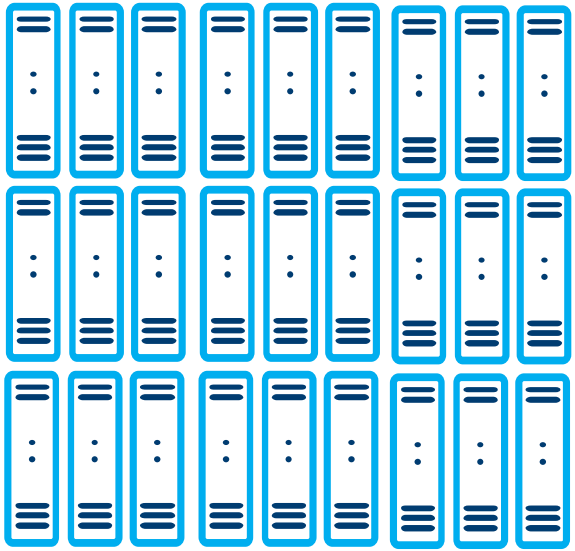
最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

HPC & AI の収束

ソフトウェア・スタック

HPC

AI

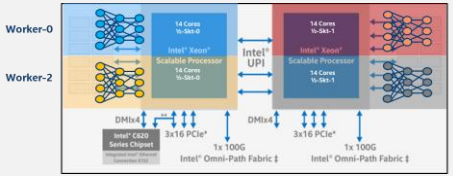


最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

MLPerf

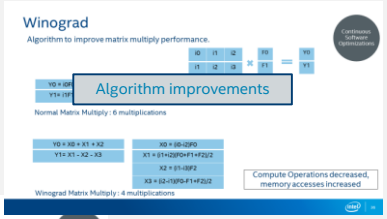
インテル® Xeon® プラットフォームにおけるディープラーニング性能の向上

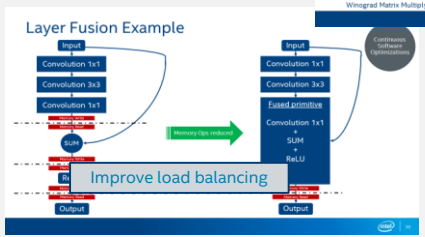
システムレベルの最適化

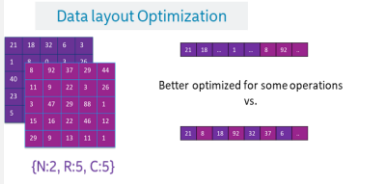


1 つのノードでフレームワークの複数のインスタンスを走らせることにより CPU 使用率を上げる

ソフトウェアの最適化





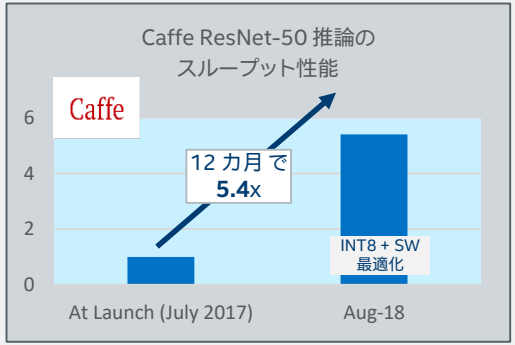


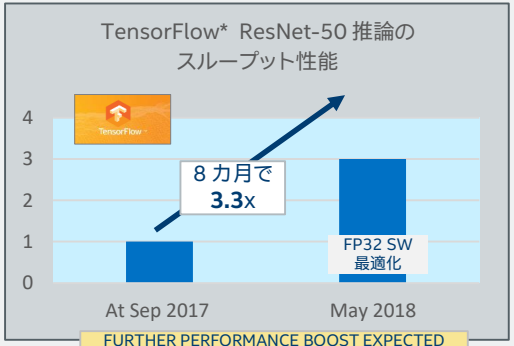
最適化されたソフトウェアで性能向上

最適化後フレームワーク vs
最適化前のフレームワークの性能

14x ResNet-50 — 訓練のスループットが改善

3.2x Inception-V3 — 推論のスループットが改善





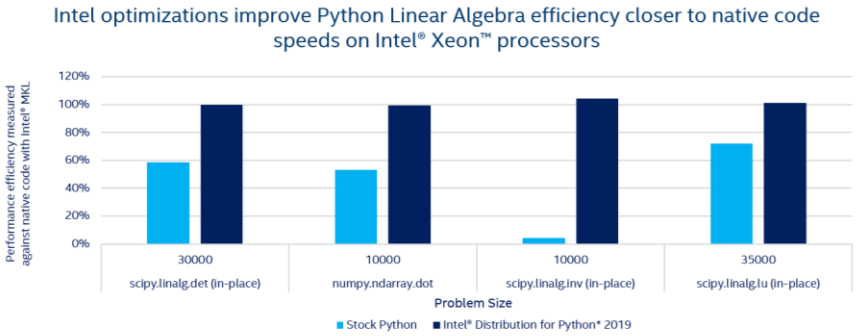
FURTHER PERFORMANCE BOOST EXPECTED WITH INT8 OPTIMIZATIONS

最適化に関する注意事項

© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® Distribution for Python* 2019

インテル® Distribution for Python* 2019 による Python* の性能向上



Performance results are based on testing as of **July 9, 2018** and may not reflect all publicly available security updates. See configuration disclosure for details. No product can be absolutely secure.
Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information, see Performance Benchmark Test Disclosure.

Testing by Intel as of July 9, 2018. Configuration: Stock Python: python 3.6.6 hc3d631a_0 installed from conda, numpy 1.15, numba 0.39.0, llvmlite 0.24.0, scipy 1.1.0, scikit-learn 0.19.2 installed from pip; Intel Python: Intel® Distribution for Python* 2019 Gold: python 3.6.5 intel_11, numpy 1.14.3 intel_py36_5, mkl 2019.0 intel_101, mkl_fft 1.0.2 intel_np114py36_6,mkl_random 1.0.1 intel_np114py36_6, numba 0.39.0 intel_np114py36_0, llvmlite 0.24.0 intel_py36_0, scipy 1.1.0 intel_np114py36_6, scikit-learn 0.19.1 intel_np114py36_35; OS: CentOS Linux 7.3.1611, kernel 3.10.0-514.el7.x86_64; Hardware: Intel(R) Xeon(R) Gold 6140 CPU @ 2.30GHz (2 sockets, 18 cores/socket, HT:off), 256 GB of DDR4 RAM, 16 DIMMs of 16 GB@2666MHz

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice revision #20110804.
For more complete information about compiler optimizations, see our Optimization Notice.

インテル® Distribution for Python* の線形代数関数は
従来の Python* 関数よりも高速

高性能の Python* ディストリビューション

- 高速な NumPy*、SciPy*、scikit-learn は科学技術計算、マシンラーニング、データ解析に最適
- 使用している Python* にドロップインで置き換え、コードの変更は必要ない
- 最新のインテル® プロセッサー向けに高度に最適化
- プライオリティー・サポート (英語) の利用 – 技術的な問題はインテルのエンジニアに直接質問[1]

新機能

- scikit-learn でより高速のマシンラーニング: サポート・ベクトル・マシン (SVM) と K 平均法予測はインテル® DAAL で高速化
- インテル® Parallel Studio XE 2019 インストーラーに統合、簡単なコマンドラインのスタンドアロン・インストール可能
- XGBoost パッケージの同梱 (Linux* のみ)

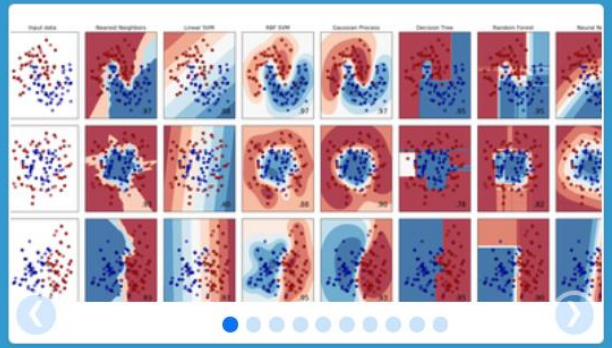
[1] 有償バージョンのみ

性能に関するテストに使用されるソフトウェアとワークロードは、性能がインテル® マイクロプロセッサー用に最適化されていることがあります。SYSmark* や MobileMark* などの性能テストは、特定のコンピューター・システム、コンポーネント、ソフトウェア、操作、機能に基づいて行ったものです。結果はこれらの要因によって異なります。製品の購入を検討される場合は、他の製品と組み合わせた場合の本製品の性能など、ほかの情報や性能テストも参考にして、パフォーマンスを総合的に評価することをお勧めします。詳細については、http://www.intel.com/performance (英語) を参照してください。

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

最適化に関する注意事項: インテル® コンパイラーでは、インテル® マイクロプロセッサーに限定されない最適化に関して、他社製マイクロプロセッサー用に同等の最適化を行えないことがあります。これには、インテル® ストリーミング SIMD 拡張命令 2、インテル® ストリーミング SIMD 拡張命令 3、インテル® ストリーミング SIMD 拡張命令 3 補足命令などの最適化が該当します。インテルは、他社製マイクロプロセッサーに関して、いかなる最適化の利用、機能、または効果も保証いたしません。本製品のマイクロプロセッサー依存の最適化は、インテル® マイクロプロセッサーでの使用を前提としています。インテル® マイクロアーキテクチャーに限定されない最適化のなかにも、インテル® マイクロプロセッサー用のものがあります。この注意事項で言及した命令セットの詳細については、該当する製品のユーザー・リファレンス・ガイドを参照してください。注意事項の改訂 #20110804

Python* 向けの最も使われるマシンラーニング・パッケージ



scikit-learn

Machine Learning in Python

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

Classification

Identifying to which category an object belongs to.

Applications: Spam detection, Image recognition.
Algorithms: SVM, nearest neighbors, random forest, …

— Examples

Regression

Predicting a continuous-valued attribute associated with an object.

Applications: Drug response, Stock prices.
Algorithms: SVR, ridge regression, Lasso, …

— Examples

Clustering

Automatic grouping of similar objects into sets.

Applications: Customer segmentation, Grouping experiment outcomes
Algorithms: k-Means, spectral clustering, mean-shift, …

— Examples

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® DAAL で scikit-learn の最適化

scikit-learn

- マシンラーニングで最も使用されるパッケージ
- 異なるパラメーターで数百ものアルゴリズム
- とても融通が利いて、使いやすいインターフェイス

DAAL4Py

インテル® DAAL は Python* API (ミドルウェア) を備える

インテル® DAAL

インテル® アーキテクチャーにおける高性能な分析とマシンラーニング・アルゴリズム

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® Distribution for Python* 2019 のインストール

スタンドアロンのインストーラー	フル・インストーラーのダウンロード https://software.intel.com/en-us/intel-distribution-for-python (英語)
Anaconda.org Anaconda.org/intel channel	> conda config --add channels intel > conda install intelpython3_full > conda install intelpython3_core
PyPI	> pip install intel-numpy > pip install intel-scipy + インテル® ライブラリーのランタイムパッケージ > pip install mkl_fft + インテルの開発パッケージ > pip install mkl_random
Docker* Hub	docker pull intelpython/intelpython3_full
YUM/APT	yum/apt でアクセス https://software.intel.com/en-us/articles/installing-intel-free-libs-and-python (英語)

2.7 & 3.6
(3.7 coming soon)

Linux*
Windows*
macOS*

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

Python* + ネイティブコードのチューニング

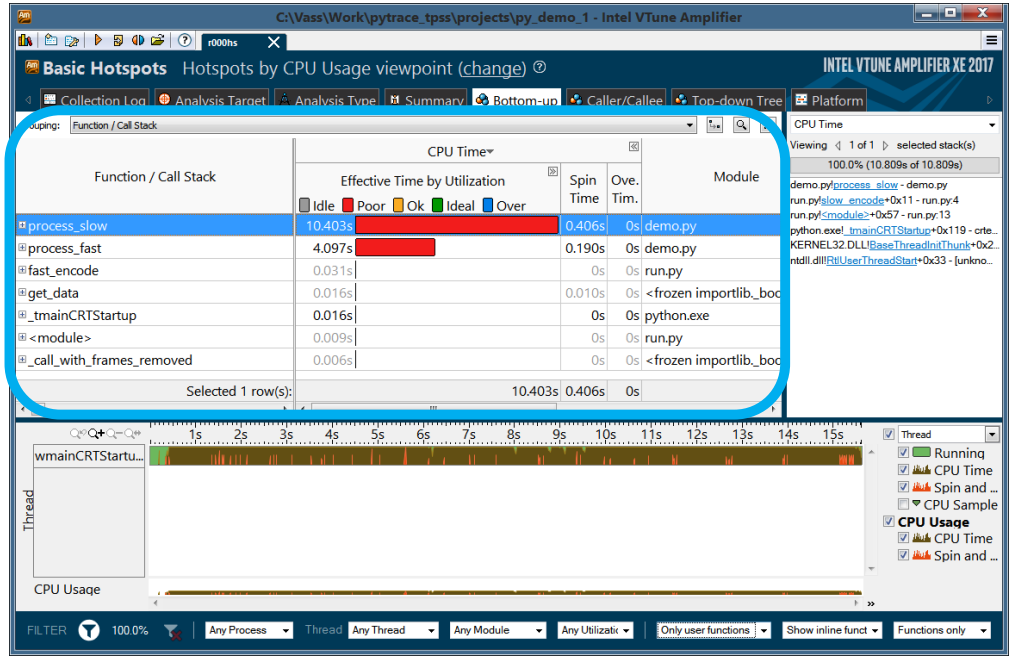
インテル® VTune™ Amplifier で性能解析 (インテル® Parallel Studio XE に同梱)

課題

- Python* とネイティブコードが混在するアプリケーションを 1 つのツールでプロファイル
- 非効率なランタイム実行を検出

ソリューション

- Python*/C/C++ 混在コードと拡張を自動検出
- パフォーマンス hotspot を行レベルで正確に識別
- 低オーバーヘッド、実行中のアプリケーションへのアタッチ / デタッチ
- パフォーマンスへの影響が大きい箇所をチューニング



Python* とネイティブ関数の自動検出およびパフォーマンス解析

インテル® VTune™ Amplifier とインテル® Parallel Studio XE で可能

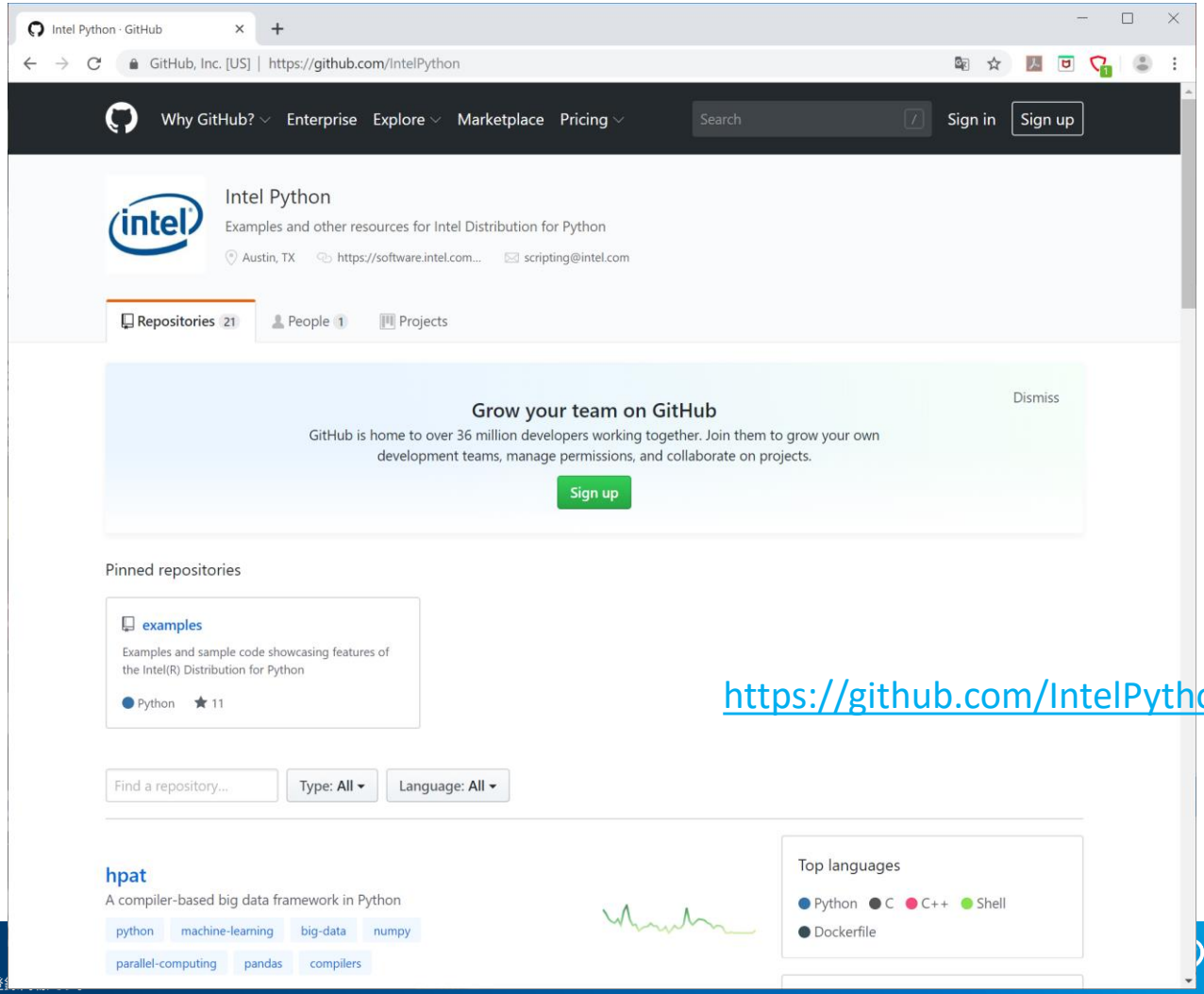
最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

IDP デモ

Numpy/random

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® Distribution for Python* のリソース



https://github.com/IntelPython

サンプルプログラムやベンチマーク等

最適化に関する注意事項

© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録

インテル® パフォーマンス・ライブラリー

インテル® マス・カーネル・ライブラリー (インテル® MKL)
インテル® データ・アナリティクス・アクセラレーション・ライブラリー (インテル® DAAL)

インテル® マス・カーネル・ライブラリー

(インテル® MKL)

インテル® MKL の機能

線形代数
- BLAS
- LAPACK
- ScaLAPACK
- スパース BLAS
- 反復法スパース ソルバー
- PARDISO
- クラスター・スパース・ソルバー

FFT
- 多次元
- FFTW インターフェイス
- クラスター FFT

ベクトル RNG
- 合同数
- Wichmann-Hill
- メルセンヌ ツイスター
- Sobol
- Neiderreiter
- 非決定的

サマリー統計
- 尖度
- 変化係数
- 順序統計量
- 最小/最大
- 分散/共分散

ベクトル演算
- 三角
- 双曲線
- 指数
- 対数
- 累乗
- 累乗根

その他
- スプライン
- 補間
- 信頼領域
- 高速ポアソン ソルバー

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® データ・アナリティクス・アクセラレーション・ライブラリー

(インテル® DAAL)

インテル® DAAL によるエッジでの解析のスピードアップ

マシンラーニングおよびビッグデータ解析のパフォーマンスを高速化

- アプリケーションがより速くより優れた予測を行えるように支援
- 最高のパフォーマンスを引き出せるようにデータの取り込みとアルゴリズムの計算を最適化
- 広範なアプリケーションのニーズを満たす、オフライン、ストリーミング、分散型使用モデルをサポート
- エッジデバイスとクラウドサービス間で解析ワークロードを分割してアプリケーション全体のスループットを最適化

詳細: isus.jp/intel-daal/

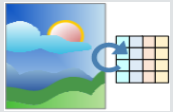
バージョン 2019 の新機能

新しいアルゴリズム

- **ハイパフォーマンス・ロジスティック回帰**、最も広く利用されている分類アルゴリズム
- **拡張勾配ブースティング機能**、ユーザー定義のコールバックによる不正確な分割計算およびアルゴリズム・レベルの計算のキャンセルにより柔軟性を拡張
- **ユーザー定義の CSV/IDBC データソースのデータ変更プロシージャー**、広範な特徴抽出および変換手法を実装

前処理 → 変換 → 解析 → モデリング → 検証 → 意思決定

展開、
フィルタリング、
正規化

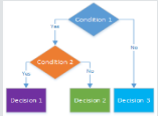


集計、
次元縮小

サマリー統計、
クラスタリングなど

マシンラーニング (訓練)、
パラメーター推定、
シミュレーション

仮説検証、
モデルエラー

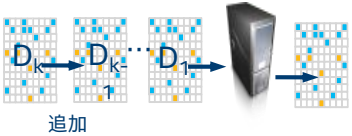


予測、
決定木など

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

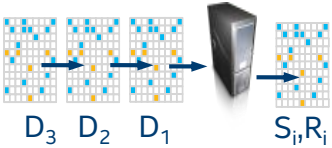
処理モデル

バッチ処理



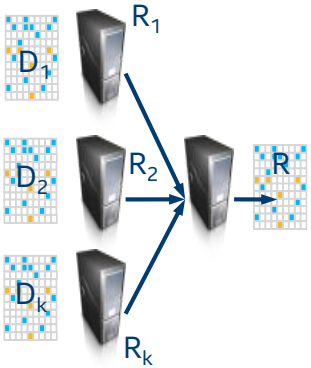
$R = F(D_1,\cdots,D_k)$

オンライン処理



$S_{i+1} = T(S_i,D_i)$
$R_{i+1} = F(S_{i+1})$

分散処理



$R = F(R_1,\cdots,R_k)$

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

データ変換および解析アルゴリズム

インテル® DAAL

- データセットの基礎統計
  - 低次モーメント
  - 分位数
  - 順序統計量
- 相関と依存関係
  - コサイン距離
  - 相関距離
  - 分散共分散行列
- 行列の因数分解
  - SVD
  - QR
  - コレスキー
- 次元縮小
  - PCA
  - 相関ルールマイニング(アプリオリ)
  - 最適化ソルバー(SGD、AdaGrad、lBFGS)
- 外れ値検出
  - 単変量
  - 多変量
  - 数学関数(exp、log、…)

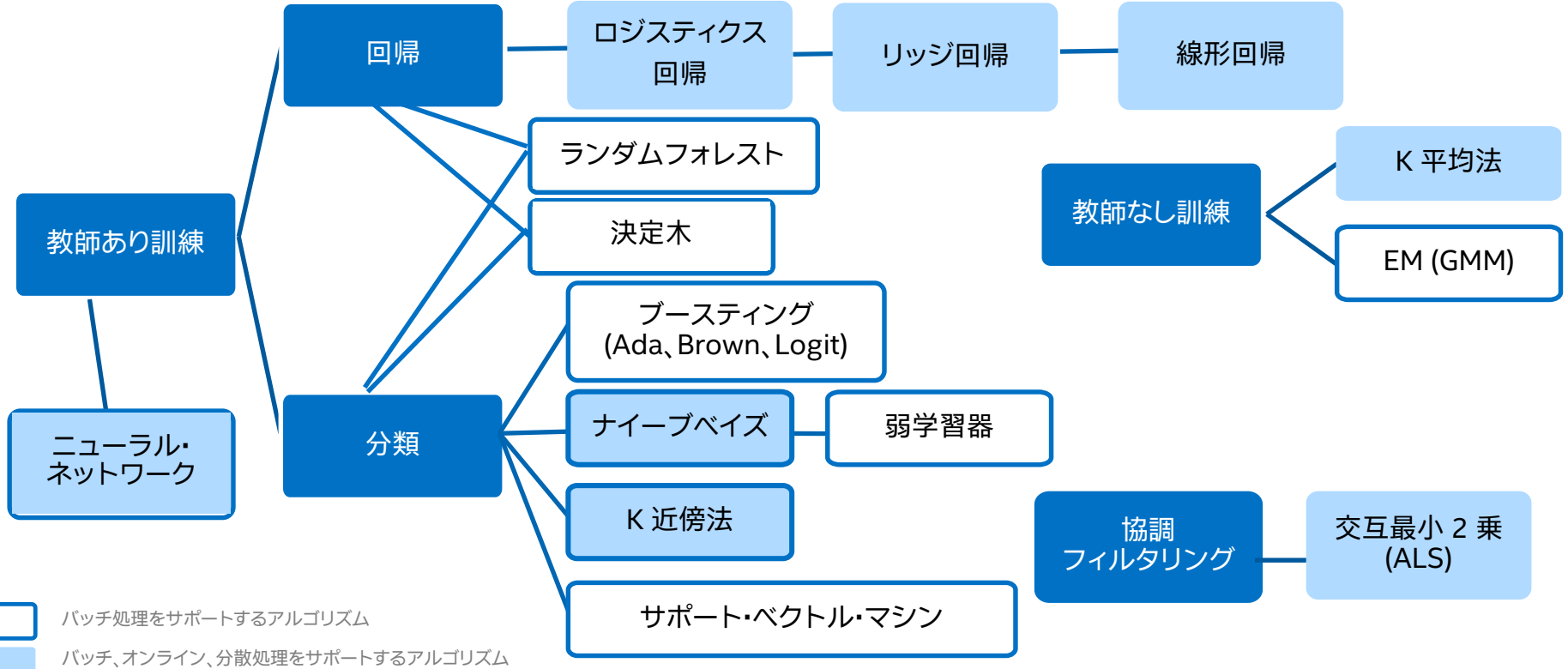
バッチ処理をサポートするアルゴリズム

バッチ、オンライン、分散処理をサポートするアルゴリズム

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

マシンラーニング・アルゴリズム

インテル® DAAL



バッチ処理をサポートするアルゴリズム

バッチ、オンライン、分散処理をサポートするアルゴリズム

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

インテル® DAAL による性能のスケール

CPU コア内

- SIMD ベクトル化: 最新の命令セットに最適化、インテル® AVX2、インテル® AVX-512...
- 内部的には、インテル® MKL のシーケンシャル部分を採用

マルチコアやメニーコアへのスケール

- スレッディング・ビルディング・ブロックによるスレッド化

クラスターへのスケール

- ユーザーのアプリケーションによる分散処理 (MPI、MapReduce など)
- インテル® DAAL の利用
  - 部分的で中間的な計算結果をサポートするデータ構造
  - 部分的で中間的な計算結果を全体の計算結果にまとめる関数

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

DAAL4Py 分散処理のデモ

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

サマリー/Q&A

さあ、始めよう

最適化に役立つ情報

iSUS (IA Software User Society)

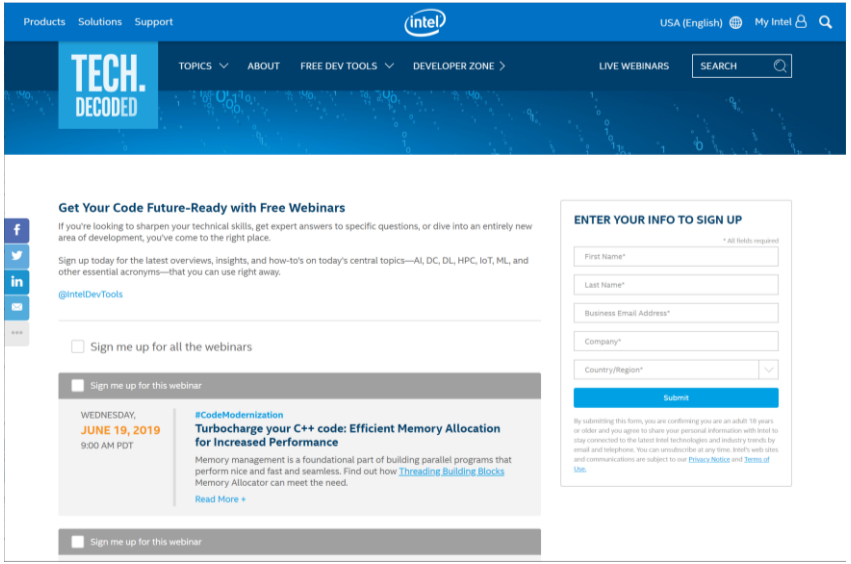
- インテル® ソフトウェア開発製品の日本語版や日本語ヘルプの提供
- インテル コーポレーションのインテル® ソフトウェア・ネットワークの記事の日本語化

インテル® Parallel Studio XE

- 概要、機能、サポート、コードサンプル (英語)
- トレーニング資料、Tech.Decoded ウェビナー (英語)、HOW TO ビデオと記事 (英語)
- お客様の声 (英語) & ケーススタディー
- その他のインテル® ソフトウェア開発製品

インテル® Code Modernization Program

- 概要 (英語)
- ライブ・トレーニング (英語)



TECH ウェビナーの視聴　https://intel.ly/2PdkNhN (英語)

最適化に関する注意事項
© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

法務上の注意書きと最適化に関する注意事項

パフォーマンス結果は 2018 年 9 月時点のテスト結果に基づいたものであり、公開されている利用可能なすべてのセキュリティー・アップデートが適用されていない可能性があります。詳細については、構成の開示を参照してください。絶対的なセキュリティーを提供できる製品はありません。

性能に関するテストに使用されるソフトウェアとワークロードは、性能がインテル® マイクロプロセッサー用に最適化されていることがあります。SYSmark* や MobileMark* などの性能テストは、特定のコンピューター・システム、コンポーネント、ソフトウェア、操作、機能に基づいて行ったものです。結果はこれらの要因によって異なります。製品の購入を検討される場合は、他の製品と組み合わせた場合の本製品の性能など、ほかの情報や性能テストも参考にして、パフォーマンスを総合的に評価することをお勧めします。詳細については、www.intel.com/benchmarks (英語) を参照してください。

本資料の情報は、現状のまま提供され、本資料は、明示されているか否かにかかわらず、また禁反言によるとよらずにかかわらず、いかなる知的財産権のライセンスも許諾するものではありません。製品に付属の売買契約書『Intel's Terms and Conditions of Sale』に規定されている場合を除き、インテルはいかなる責任を負うものではなく、またインテル製品の販売や使用に関する明示または黙示の保証 (特定目的への適合性、商品性に関する保証、第三者の特許権、著作権、その他、知的財産権の侵害への保証を含む) をするものではありません。

© 2019 Intel Corporation. 無断での引用、転載を禁じます。Intel、インテル、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Arria、Intel Atom、Intel Core、Iris、Stratix、Xeon、Movidius、Myriad、Intel Nervana、OpenVINO は、アメリカ合衆国および/ またはその他の国における Intel Corporation またはその子会社の商標です。

最適化に関する注意事項

インテル® コンパイラーでは、インテル® マイクロプロセッサーに限定されない最適化に関して、他社製マイクロプロセッサー用に同等の最適化を行えないことがあります。これには、インテル® ストリーミング SIMD 拡張命令 2、インテル® ストリーミング SIMD 拡張命令 3、インテル® ストリーミング SIMD 拡張命令 3 補足命令などの最適化が該当します。インテルは、他社製マイクロプロセッサーに関して、いかなる最適化の利用、機能、または効果も保証いたしません。本製品のマイクロプロセッサー依存の最適化は、インテル® マイクロプロセッサーでの使用を前提としています。インテル® マイクロアーキテクチャーに限定されない最適化のなかにも、インテル® マイクロプロセッサー用のものがあります。この注意事項で言及した命令セットの詳細については、該当する製品のユーザー・リファレンス・ガイドを参照してください。
注意事項の改訂 #20110804

© 2019 Intel Corporation. 無断での引用、転載を禁じます。
* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。