

### 人工知能(AI)を実現するディープラーニング(深層学習)の 統合開発環境「Neural Network Console クラウド版」にて 膨大な計算を可能にする分散学習環境の提供を開始

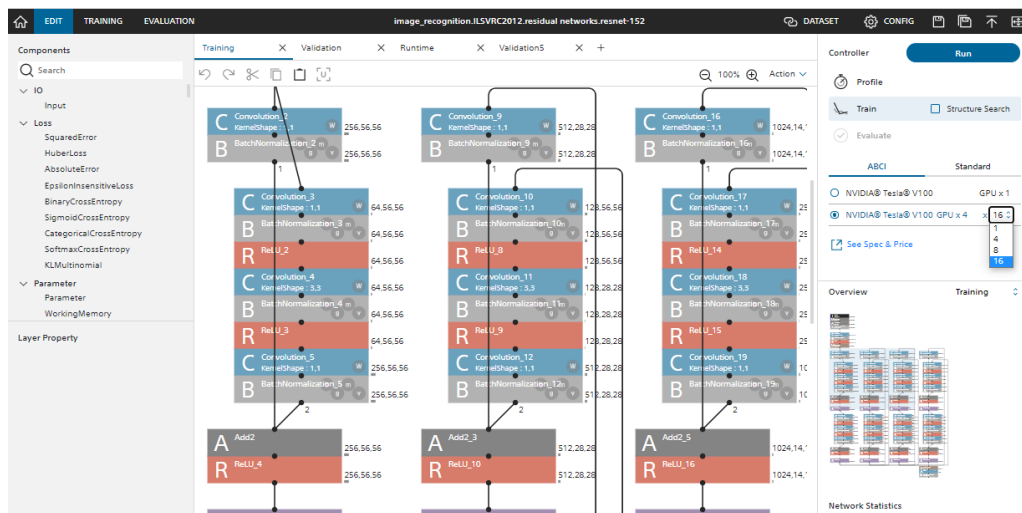
ソニーは、ディープラーニング(深層学習)のプログラムを生成できる統合開発環境「コンソールソフトウェア: Neural Network Console クラウド版」(<https://dl.sony.com/ja/>)において、複数のノードを利用して膨大な計算を可能にする分散学習環境の提供を、本日より開始します。

## Neural Network Console

ソニーは、2017年6月にディープラーニング開発のためのコアライブラリ「Neural Network Libraries (<https://nnabla.org/>)」をオープンソース化し、同年8月にコンソールソフトウェア「Neural Network Console」の提供を開始しました。また、2018年5月からは複数GPUによる高速学習サービス「Neural Network Console クラウド版」の提供を開始し、ウェブブラウザでアクセスするだけで、GUIベースの直観的な操作画面やクラウド上のリソースを使用した本格的なディープラーニングのプログラム開発ができる統合開発環境を提供しています。

さらに、2019年4月よりクラウド版のGPUとして、国立研究開発法人 産業技術総合研究所(以下:産総研)が構築・運用する世界最大規模のAI処理向け計算インフラストラクチャであるAI橋渡しクラウド(AI Bridging Cloud Infrastructure、以下「ABCI」)の選択ができるようになりました。

このたび、「Neural Network Console クラウド版」の学習・評価で利用できる「ABCI」の計算ノード(※1)において、複数のノード(マルチノード)による分散学習に対応します。これまで1つの学習処理あたりで利用できる計算ノードは1台(4GPU)が最大でしたが、今後は最大16台(64GPU)の利用が可能になります。これにより、従来は一部の研究者などが利用してきた大規模なAI開発環境を、一般のユーザーがGUIを介して手軽に利用できるようになります。



ディープラーニングとは、人間の脳を模倣したニューラルネットワークを用いた機械学習の一手法です。これを使用することで画像認識や音声認識の性能が近年飛躍的に向上しています。

一方で、認識精度を向上させるために学習データのサイズやモデルのパラメーター数が増加し、計算時間が飛躍的に増加して、一度の学習に数週間から数カ月を要するケースもでてきました。AI 開発においては様々な試行錯誤を繰り返す必要があるため、学習時間の短縮は非常に重要となります。そのため、複数の GPU を活用した分散学習による開発の効率化が現在脚光を浴びています。

ソニーは、産総研が 2018 年 10 月に実施した「ABCI グランドチャレンジ」にて当時世界最速(※2)となる学習速度を達成するなど、かねてより分散学習の研究開発に取り組んできました。これらの知見をもとに、学習高速化の様々な技術をユーザーが使いやすい GUI の形にし、誰もが大規模な分散学習を実現できるサービスとして提供するに至りました。

AI による利便性の向上が多く製品やサービスにおいて期待されるなか、研究者や開発者、事業者など幅広いユーザーの皆様が、高度なプログラミングをより効率的に実現できるよう今後もサポートしてまいります。

## ○ 「Neural Network Console クラウド版」マルチノード分散学習の主な特長

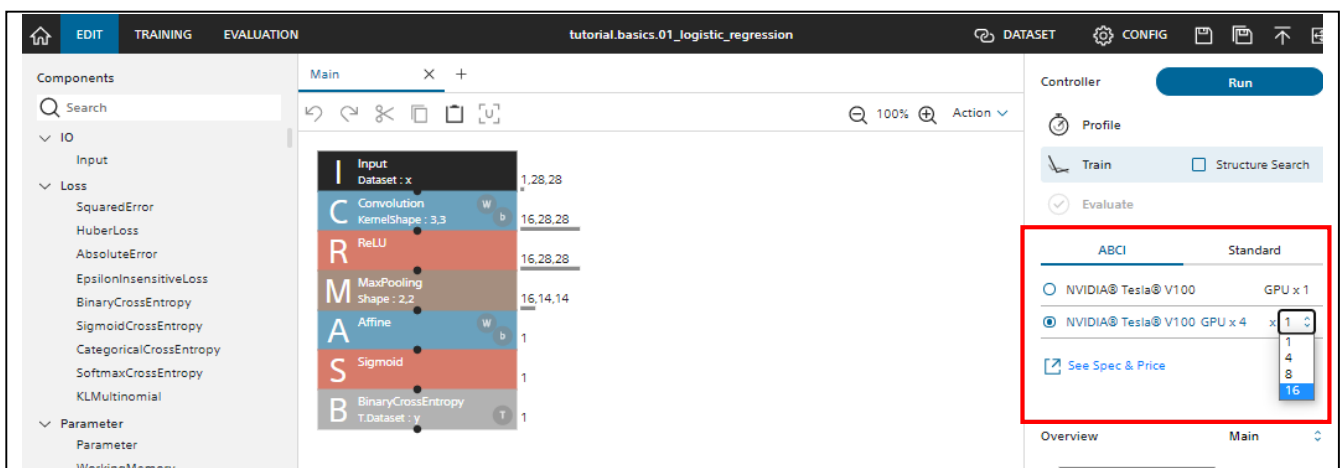
### “GUI×分散学習”で大規模な AI 開発環境を実現

直感的でわかりやすい GUI を通じたニューラルネットワーク設計ができるディープラーニング開発環境において、新たにマルチノードによる分散学習の実行に対応しました。これまで一部の研究者などが利用していた高度な計算環境を、一般のユーザーが自身の PC からウェブインターフェースを介して、手軽に操作できるようになります。

### 複数ノードの割り当てが手軽にできるユーザーインターフェース

ノード数を選択しボタンをクリックするだけで分散学習の利用が可能です。

大量のデータを用いた複雑な学習を実行する場合も、内容に応じてノード数を選択することで柔軟に処理性能の上限をコントロールできます。オンデマンドでクラウド環境のリソースを利用することで、一時的な計算環境構築のための費用削減とともに、柔軟な学習環境での効率的な開発が可能です。



### 最大 16 の計算ノードを利用可能、演算処理が飛躍的に速く

ノード数は、4 台、8 台、16 台の3種類から選択できます。

「Neural Network Console クラウド版」では、これまで1つの学習処理あたりで利用できる計算ノードは 1 台が最大でしたが、今回、最大 16 台の利用が可能となり、演算処理が飛躍的に速くなります。

なお、今後最大ノード数を更に向上できるよう研究開発を継続してまいります。

○ 利用料金

GPU	GPU 相当数	メモリ サイズ	ノード数	料金 (1 時間)	備考
NVIDIA® TESLA® V100 GPU	GPU × 1	60GB	1	300 円	ABCI の「rt_G.small」ノードを 1 台利用
	GPU × 4	240GB	1	1,650 円	ABCI の「rt_G.large」ノードを 1 台利用
		360GB	4	6,600 円	ABCI の「rt_F」ノードを 4 台利用
			8	13,200 円	ABCI の「rt_F」ノードを 8 台利用
			16	26,400 円	ABCI の「rt_F」ノードを 16 台利用

※右下の太枠内が今回新たに提供するメニューです。

※記載の金額はすべて税抜です。

<学習済みモデルの API 公開機能追加について>

また、今回「Neural Network Console」で作成した学習済みモデルを API で公開する機能を追加しました。これまで学習済みモデルの実行においては利用者が実行環境を準備する必要がありましたが、新機能により、ユーザーが実行環境を構築・運用する必要がなくなり、クラウド上のモデルを手軽に実行することが可能となります。

※詳細は、サービスページを参照ください。

<https://support.dl.sony.com/docs-ja/モデル API 公開機能/>

(※1) AI 開発用に構成された計算システム

(※2) 2018 年 11 月 13 日時点(ソニー調べ)

ImageNet データセットを利用して ResNet-50 モデルの学習を実施しました。

ImageNet/ ResNet-50 とは、それぞれ一般的に利用されている画像認識用データセットおよび画像認識用モデルとなります。