

メタデータ、機械学習実用化フレームワーク“Montage”で 学習済ディープラーニングの生産性と品質を向上

～自動評価や特徴量の事前抽出により複数方式間の優劣判定を
効率化し、人工知能導入を加速

メタデータ株式会社（所在地：東京都文京区 代表取締役社長：野村直之）は、様々なディープラーニングの方式や他の機械学習による文章、音声、画像等の認識分類精度の評価を効率化し、様々な産業応用の局面で人工知能の導入を加速する機械学習実用化フレームワーク“Montage”を開発し導入しました。“Montage”により、学習済の専門画像認識システム等を短時日で低コストで作成し、精度報告書付で出荷出来るようになります。“Montage”の概要は、2016年7月1日午前、「コンテンツ東京内 AI・人工知能ワールドのセミナーにて講演いたします。

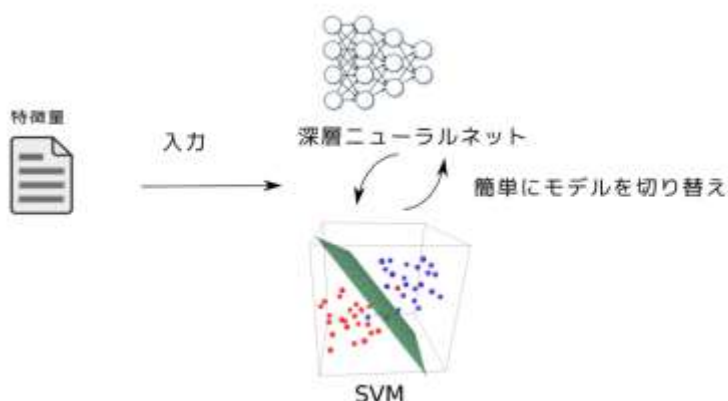
■背景

ビッグデータが溢れ、その活用のための分析が人手では追い付かず、人工知能的なソフトウェアを業務に組み込む問題意識が高まっています。特にディープラーニングが注目を浴びています。ディープラーニングが、人間がモデルを与えることなく勝手に特徴を抽出して、認識能力、分類能力を備えてくれることは画期的でした。しかし、その精度がデータの量や質に大きく依存し、本番さながらの実験を行ってみるまでは、実用性が不明という欠点があります。ディープラーニングに学習させた結果、ネットワークの内部がどのようになっているか、当面原理的に解析できない状況が続くそうなので、試行錯誤、評価が必要です。

成功／失敗、改善／改悪の原因分析が困難で追加投資、保守予算が見えず、また IT ベンダーとユーザ間の責任分界点が不明確な点が、特に、機密データを IT ベンダーに開示できない企業にとっては致命的な問題点となり得ます。

■機械学習実用化フレームワーク“Montage”の概要

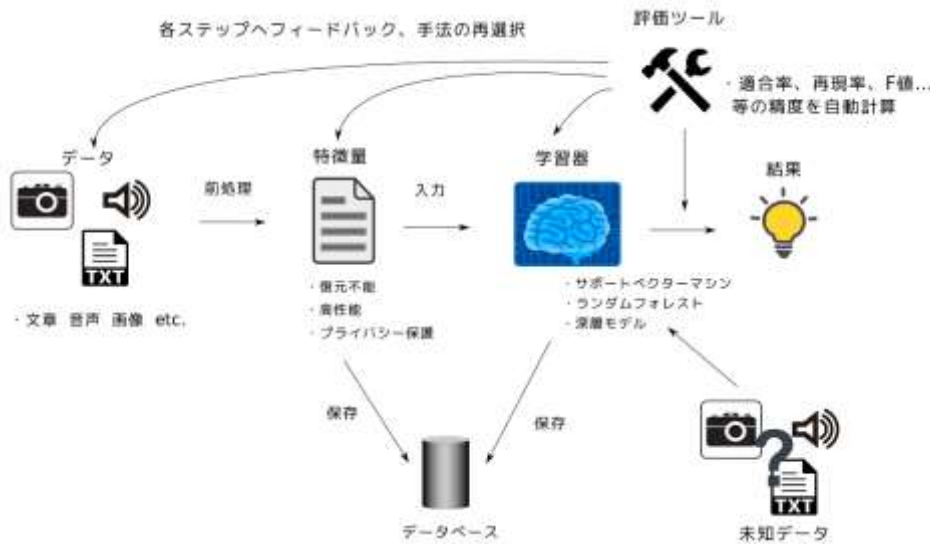
これらの問題を解決するため、メタデータ株式会社は、機械学習実用化フレームワーク“Montage”を開発しました。“Montage”のコア部分では、予め抽出しておいた特徴量群を、様々なディープラーニング（別名「深層ニューラルネット」）や、SVM (Support Vector Machine) など複数種類の学習器の間で切り替えて、認識や分類を行わせることができます。このため、



主な優れた機械学習で必要となる特徴量を高々10種類ほど、学習データから抽出しておけば、学習データを廃棄、返却して行うことができます。これは、データの所有権、著作権の問題を回避するため、長期間貸与を受けることなく、高々最初の1週間程度、権利者から利用させてもらうことで、実用精度の認識・分類エンジンを、

その複数の特徴量抽出結果(元データより何桁も情報量が少なく元データの ” 表現 ” と同等の著作権は消失していると考えられます) を使いながら鍛え上げ、チューニングしていくことで実用精度を達成することが可能です。特に、金融系の応用現場など、学習のための元データを AI 構築を担う IT ベンダーにさえ一切見せられないような場合、“Montage” の前処理、すなわち、特徴量抽出モジュールを応用現場に貸与し、マニュアル通りに特徴量抽出を行ってもらって受け取ることで、機密データを IT ベンダーが見ないまま、その種のデータから認識・分類器のトレーニングを進めたり、対話の精度を上げたりすることが可能となります。

機械学習実用化フレームワーク “Montage” のもう 1 つの大きな特徴は、精度の自動評価です。予め、クラウドソーシング等で高品質な正解データを作成しておけば、ある段階までトレーニングが進んだ学習器群の精度 (適合率、再現率、F 値等) を自動計算し、自動比較することが



ことができます。これにより、最も精度の高い学習器を選んで、さらなる追加トレーニングを施す、といった意思決定が容易になります。トータルプロセスとして、人間は人間ならではの判断による正解データ作りと、誤りの質の評価などに専念出来るようになり、AI の応用現場への投入を加速することができます。

機械学習実用化フレームワーク “Montage” の概要は、2016 年 7 月 1 日午前、「コンテンツ東京内 AI・人工知能ワールドのリアルワールド社セミナーにて講演いたします。

https://www.ct-next.jp/Conference_Event/seminar-event06/

94 メタデータ株式会社について

メタデータ株式会社は、2005 年 12 月にメタデータ活用技術、セマンティック技術の応用ソフトウェア会社として設立。適合型自然言語解析エンジン『メタパーザ』を駆使した意味解析技術により、5W1H 抽出によるコンテンツ連携や個人情報自動匿名化、評判分析(ネガポジ・感情解析)、意味による自動分類やポジショニングマップ描画、Web 対話ロボットなど、SaaS、クラウド向けアプリケーションを開発してまいりました。最近では、機械学習、ディープラーニング等の人工知能研究の蓄積を活かした分析エンジンやマッチングエンジンを構築し、画像認識・分類のための学習済ディープラーニングの販売事業や、超 AI とも呼ぶべき超高速マッチングエンジン xTech の販売にも注力しています。

なお、代表取締役・野村は、MIT マサチューセッツ工科大学・人工知能研究所の研究員時代に概念体系 WordNet の理論と活用研究に貢献。その後の、WordNet をベースに 1000 数百万枚の写真に体系的な意味ラベルを付与した ImageNet の完成によるディープラーニングの劇的精度向上にも、間接的な貢献を果たしています。

94 本件に関するお問い合わせ先

メタデータ株式会社

担当：野村 nomuran@metadata.co.jp

TEL:03-3813-5447

FAX:03-3813-5447 <http://www.metadata.co.jp>

