

特集を読む前に知っておきたい…機械学習とディープ・ラーニングの違い

新井 正敏

本特集ではディープ・ラーニングのからくりを解き明かしていきます。そもそもディープ・ラーニングとは何なのでしょう。ディープ・ラーニングは機械学習の一部に位置付けられていますが、従来の機械学習よりも学習の精度や自動化が進んでいると言われています。

そこで本稿では、従来の機械学習との違いから、ディープ・ラーニングとは何なのかについて解説します。(編集部)

本稿では、特集の内容に入る前に、ディープ・ラーニングとは何か、混乱しがちな機械学習とディープ・ラーニングの違いを示します。

そもそも…何ができるの？

● ご利益はどちらも同じ…分類と回帰

機械学習とディープ・ラーニングのご利益は大きく分けて分類(Classification)と回帰(Regression)があります。

- 分類の例：画像を入力して、その画像に写っているのが犬か猫かを予測
- 回帰の例：過去の株価を入力して、今後の連続的な株価を予測

このように、機械学習やディープ・ラーニングにおいては、入力データに対する予測を出力することになります。本稿では、分類と回帰の特徴を示した上で、機械学習とディープ・ラーニングの違いを具体的に示します。

• 分類

- 目的：複数の入力データがどの離散的なカテゴリ(ラベル)に属するかを予測する
- 予測：あらかじめ決められた離散的なラベル(例えば犬、猫、サル、鳥)から1つを予測(選択)する
- アルゴリズム：ロジスティック回帰、サポートベクタマシン(SVM)、k-近傍法(k-NN)、ニューラルネットワーク(ディープ・ラーニング)など

• 回帰

- 目的：複数の入力データに対して、連続的な数値を予測する
- 出力：連続値(例えば日本一熱い(暑い)熊谷市の最高気温)を予測する

の最高気温)を予測する

- アルゴリズム：線形回帰、決定木回帰、ニューラルネットワーク(ディープ・ラーニング)など

分類と回帰は、機械学習やディープ・ラーニングにおける予測機能であり、それぞれ異なるタイプの予測問題に適用されます。分類は離散のカテゴリの予測、回帰は連続的な数値の予測を行います。

● 適切なアルゴリズムやモデルを選ぶことが重要

重要な点は、適切なアルゴリズムとモデルを選ぶことで、さまざまな現実の予測問題に対応することができるという点です。モデルとは、後述する特徴抽出やニューラルネットワークの構成のことです。

もし、この選択を見誤ると、筆者のように負のスパイラルに陥ってしまうことになります。

共通点と違い

画像分類を例に、機械学習とディープ・ラーニングの流れを考えます。

図1に示すのは、左にある分類画像を入力して、機械学習とディープ・ラーニングを使って離散のカテゴリ(吊り橋、列車、飛行機、車などあらかじめ決められたラベル)から、吊り橋という予測結果を出した例です。

機械学習とディープ・ラーニングの関係を図2に示します。ディープ・ラーニングは機械学習に含まれています。従って、機械学習とディープ・ラーニングは共通点を持っています。共通点と違いについて示します。

● 機械学習とディープ・ラーニングの共通点

- 分類と回帰が主な予測機能である
- 入力データからパターンや規則性の特徴を学習
- データ処理と統計学の概念に基づいている
- 学習のからくりを知るには数学の知識が必要

学習のからくりを解き明かすために、必要な数学、特に微分積分、行列演算を本特集の中で示していきます。また、学習するときのデータは、玉石混交であってはなりません。玉(学習に意味のあるデータ)の取り出し方、磨く(統計学に基づき粒をそろえる)方法も数学に基づき示していきます。

なお、機械学習とディープ・ラーニングの共通項目