

18

学習結果に対する評価基準

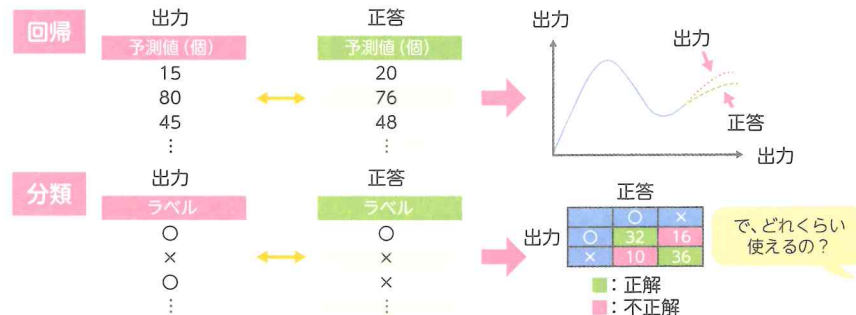
機械学習モデルの検証により、モデルの出力結果と正答の集計データを得ることができます。しかし、集計データはそのモデルの性能を端的に表現しているとは言えません。ここでは、集計データからモデルの性能を正しく表現する方法を学びましょう。

○ 機械学習の性能評価

前のセクションでは、テストデータによって機械学習モデルの性能を検証する方法を学びました。検証においてモデルにテストデータを入力すると、それぞれのデータに対し、回帰であれば何らかの数値が、分類であればデータのラベルが予測されます。それらの結果は、回帰であればそれぞれの入力において予測と正答がどれくらい離れているのかが値として現れます。一方の分類であれば、それぞれの入力に対して何のラベルに分類されたのか、表として集計できます。しかしこのような集計だけでは、「このモデルはどんな性能なのか」、あるいは「このモデルによる事業への利益は見込めるのか」といった重要な疑問に答えられません。

そこで必要なのが、テストデータを利用した検証結果から性能を「評価」することです。このセクションでは、回帰と分類においてモデルを評価する指標を紹介します。

■ 検証だけでは不十分



○ 回帰モデルの性能を評価する

回帰モデルの性能は基本的に、出力と正答の数値の差分である「予測誤差」によって評価できるため、回帰モデルの評価指標における違いは、この予測誤差をどのように集計するかの違いといえます。さっそく次のページから、代表的な R^2 (決定係数)、RMSE (平方平均二乗誤差)、MAE (平均絶対誤差) を紹介します。

■ 回帰モデルにおける評価指標

