

17

テストデータによる
予測結果の検証

機械学習アルゴリズムにおいて未知のデータに対する予測・分類性能の検証はとても重要ですが、検証の方法を間違えるとその結果はまったく意味のないものになってしまいます。ここではテストデータによる正しい検証方法を学びましょう。

○ 汎化性能とは

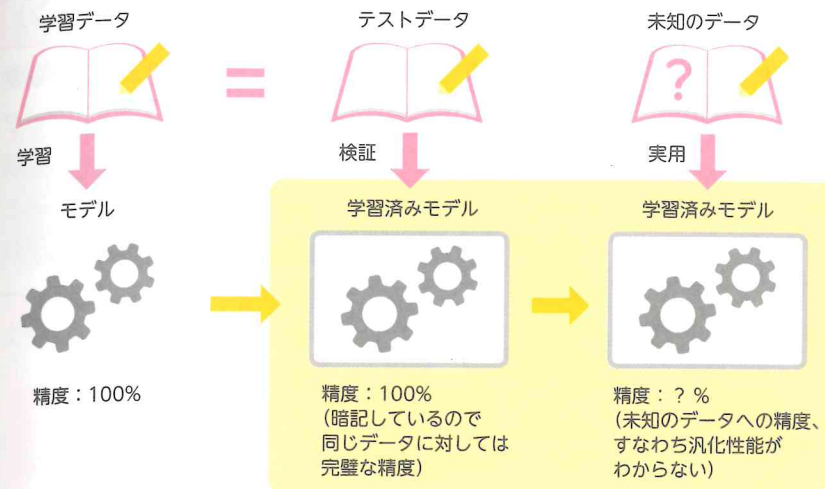
機械学習において重要なことは、データを学習することで未知のデータの予測や分類を行えるようになることです。この未知のデータに対する予測や分類の精度のことを**汎化性能**と呼びます。学習では、学習データに対する性能を基準にパラメータを更新し進めていきますが、学習が終わった段階ではまだ未知のデータに対する性能が保証されていません。そこで汎化性能を検証する必要があります。

汎化性能を検証するうえで大事なことは、「学習に利用したデータを検証に使わない」ということです。そのためには、学習データとは別に検証するためのデータを分けて残しておく必要があります。この検証用のデータを、**テスト(検証・評価)データ**と呼びます。

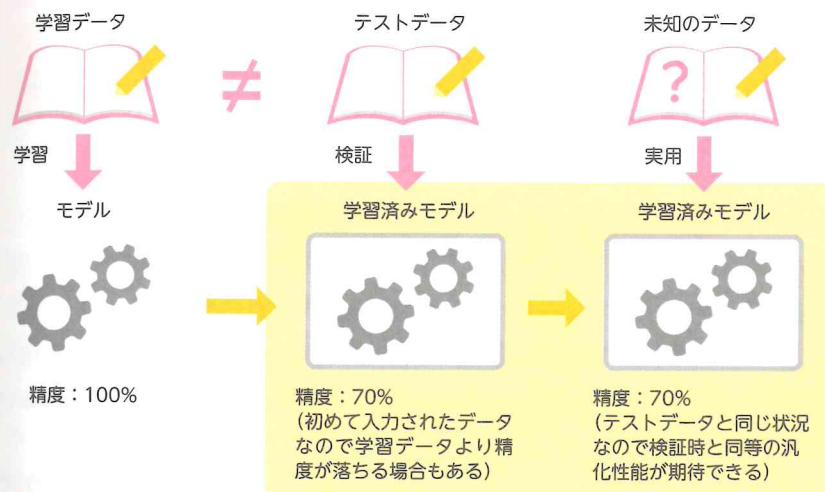
学習データとテストデータをわけることで、学習したモデルが未知のデータに対してどのくらいの性能があるのかを検証することができ、そこで初めてそのモデルの汎化性能を評価できます。作成したモデルで期待される性能がわからないことは、モデルの信頼性を保証できないということであり、アルゴリズムを実際に活用する上で大きな支障となるのです。

なお、学習データをそのままテストデータとして使用すると、検証でのモデルの精度は高くなりますが、これは汎化性能の高さを示したことはありません。勉強した問題がそっくりそのままテストに出てしまっただけは本当の学力がわからないのと同じです。

■ 学習データをテストデータとして使用すると……



■ テストデータは学習データとして使わない



それでは、実際に学習データとテストデータをどうわけばよいのでしょうか。いくつか方法がありますが、代表的な手法として、次ページからホールドアウト検証とK-分割交差検証(K-foldクロスバリデーション)を紹介します。