

ランダムフォレスト

機械学習とは、コンピュータにデータを学習させ、アルゴリズムに基づき分類や予測などのタスクを遂行する技術である。ランダムフォレストは、機械学習のアルゴリズムの1つであり、学習データの一部をランダムに抽出し、そのデータで決定木を作ることを繰り返して、複数の学習結果を作り、その結果を合議（多くの場合は、多数決）するもののことを言う（図-1）。複数の決定木を用いて結果を推定することから、1つの決定木より高精度になるため、精度が必要なケースで用いられることがある。

決定木とは、推定したい対象である目的変数の値（例えば、水を飲むまたは水を飲まない）を分ける条件を、正解を推定するのに用いる他のデータである説明変数（例えば、気温、湿度、天候など）の中から見付けるアルゴリズムのことを言う。

ランダムフォレストは、解釈性と精度のバランスがよく、実用的なアルゴリズムの1つとされている。

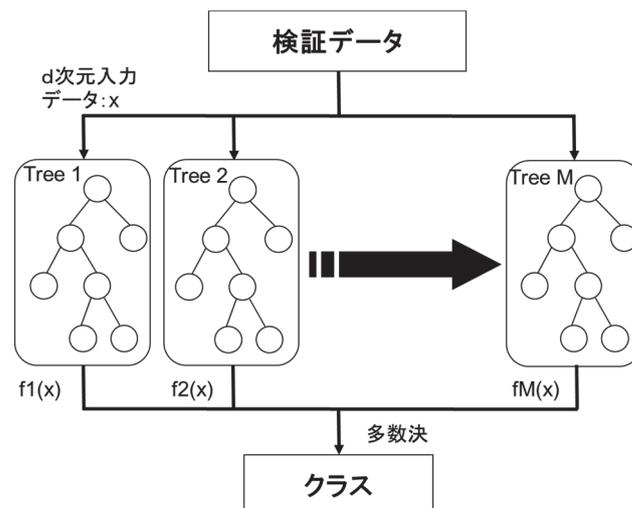


図-1 ランダムフォレストの概念図

参考文献

- 1) 本橋洋介：人工知能システムのプロジェクトがわかる本、株式会社翔泳社、pp.102～103、2018