

Regression（回帰）

モデル	Pros	Cons
線形回帰	データの量に関係なく使える。特徴量の関係についての情報を得ることができる	線形回帰の前提を満たす必要がある
多項式回帰	データの量に関係なく使え、非線形の問題に対して良い当てはまりを見せる場合がある	多項式の次数を適切に定めないと過学習してしまいがちになる
サポートベクトル回帰	適用が簡単であり、非線形の問題でも良い当てはまりを見せる。外れ値の影響を受けない	フィーチャースケーリングを行う必要がある。元のデータに対する理解が浅いと、結果を理解することが難しくなる
回帰木	フィーチャースケーリングをする必要がなく、線形/非線形のどちらにも使える	データの数が少ないと精度が低くなり、また過学習しがち
ランダムフォレスト	非線形/線形を問わず使うことができ、更に予測の精度も高い場合が多い	直観的な結果の解釈が難しい場合があり、且つ過学習しやすい。また、木の数を決める必要がある