

scikit-learn seminar

1.2.5. Tips on Practical Use

Ibaraki University

Hironori Kikuchi

Tips (1)

- Kernel cache size (for SVC, SVR, NuSVC and nuSVR)
 - 実行時間に影響を与える
 - 十分なリソースがあるならデフォルトよりも大きな値に設定する事を推奨 (500MBか1000MB)
- Setting C
 - Cはデフォルトで"1"に設定してある
 - ノイズが多い場合は値を小さく設定する

Tips(2)

- Parameter "nu" (in NuSVC, OneClassSVM and nuSVR)
 - トレーニングデータのエラー率とサポートベクターを指定する
- data for classification are unbalanced (in SVC)
 - class_weight="auto" に設定
 - ペナルティ項 "C" の値を変える
- SVM is not scale invariant
 - データを正規化してから用いる

Tips(3)

- TOL parameter
 - LinearSVCではモデルを適合する際の機能を選択する際に乱数を使用している
 - 誤差TOLのパラメータを小さくする事で結果の振れを軽減する