

Python による演習は、原則として「ML 演習番号.ipynb」ファイルが解答です。たとえば実践演習 1-1 の解答は「ML1-1.ipynb」です。

以下では、Weka を用いた演習の解答例を示します。

実践演習 3-1

iris.arff データでは、minNumObj=1, unpruned = True にすることで精度 100% が実現できます。一方、iris.2D.arff データでは 1 事例の誤りをどうしても消すことができません。これはデータ中に同じ特徴ベクトルに対して異なるクラスが正解として付与されているものがあるからです。

実践演習 3-2

minNumObj=2~4, unpruned = False にすることで精度 96% が実現できます。これに加えて useMDL-Correction（数値特徴を分割するときの基準）を False にすると、精度 96.67% が実現できます。

実践演習 3-3

minNumObj=21 にすることで精度 73% が実現できます。このときの木は、葉に近い purpose の特徴で広がっているので複雑そうに見えるだけで、全体としては比較的単純です。minNumObj=25 にすると精度 72.1% に落ちますが、木はさらに単純になります。