

課題 2: 例からの学習
(提出締切 5月22日)

目的: ニューラルネットワーク(神経回路モデル)と呼ばれているものがどういうものをコンピュータシミュレーションを通して知る. 特に, 学習が進む様子をコンピュータシミュレーションにより確認し, 速さの異なる二つ(ニューロンの活動と結合係数)のダイナミクスがあることを理解する(ソースコードを読む練習も兼ねる)

1. 「りんご」と「みかん」の分類(学習がやさしいデータ)

りんごとみかんから測定されたデータ(2次元)例が, それぞれについて10個ある. この例を参考に, りんごとみかんを分別できるような機械を学習により作りたい.

(a) データやサンプルプログラムは, 講義の web ページよりダウンロードできる.

<http://www.cs.miyazaki-u.ac.jp/~date/lectures/2006bis/>

- りんごから計測した2次元データ10個 (apples.dat)
- みかんから計測して2次元データ10個 (oranges.dat)

(b) サンプルプログラム (lab060508a.c) をコンパイルして実行する.

(c) りんごとみかんのデータがそれぞれ + と - でプロットされており, 学習が進む様子を確認する.

(d) 課題 A: 横軸に学習回数, 縦軸に正解率(またはエラー率)をプロットした図を書く.

(e) 課題 B: サンプルプログラムに書かれている学習係数などのパラメータ値を変更し, それにより学習が進む過程がどう変化するか観察し, パラメータの意味を考察する.

変更点: 学習係数 η , シグモイド関数のパラメータ λ など.

2. りんごとみかんの分類(学習が少し難しいデータ)

(a) 問題は1.と同じであるが, それぞれについてデータが20個ずつある.

ファイル名は apples3.dat と oranges3.dat .

(b) 1.と同じプログラムで試す(プログラム中のファイル名とデータの個数を変更すること).

(c) 課題 C: 横軸に学習回数, 縦軸に正解率をプロットした図を書く.

(d) 課題 D: おそらくほとんどの場合, うまく学習がすすまない. その理由を考察する.

(e) サンプルプログラム (lab060508b.c) をコンパイルして実行する. 学習が進む様子を確認する. これはバックプロパゲーション(誤差逆伝搬)と呼ばれる手法を使っている.

3. 自由課題: 自分で適当なデータを作り, どのような場合に学習が進むか, 進まないか, いろいろ試して考察する. 面白い振る舞いをする例を探す.

4. 補足: 細かい点(学習回数など)は自分で適当に決めて進めればよい. その他, 課題に曖昧な点があった場合も適当に解釈していい. その場合, どこをどう解釈したか, そう解釈した理由等をレポートに記述. レポートの最後に, 感想, 質問等を自由に記述. レポートの L^AT_EX を使った簡単な書き方は

<http://www.cs.miyazaki-u.ac.jp/~date/lectures/latex/latexreport.html> を参照.