

宿題 1

最近傍識別器によって 0-9 までの 10 クラスの手書き文字認識を行う。近傍数 k は識別誤差に関する交差確認により決定する。

以下に結果を示す。近傍数 k の候補として、1-10 を考えた。各 k についての、交差確認法で求めた正解数の平均値を表 1 に示す。これより、最も高い正解数を与える k を選ぶと、 $k = 3$ となる。 $k = 3$ に対する混同行列は表 2 のようになり、また、各カテゴリごとの正解率等は表 3 のようになった。

プログラムは??ページの Listing ??に示した。

表 1: 各 k についての、交差確認法で求めた正解数の平均値

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 484.7 | 483.2 | 485.1 | 484.8 | 483.4 | 482.7 | 481.6 | 481.2 | 480.0 | 479.7 |

表 2: 混同行列

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0 | 198 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 1 | 0 | 199 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 195 | 0 | 0 | 0 | 0 | 2 | 3 | 0 |
| 3 | 0 | 0 | 0 | 190 | 0 | 4 | 0 | 1 | 4 | 1 |
| 4 | 0 | 1 | 0 | 0 | 189 | 0 | 3 | 0 | 0 | 7 |
| 5 | 2 | 0 | 3 | 3 | 1 | 186 | 0 | 0 | 1 | 4 |
| 6 | 1 | 0 | 2 | 0 | 0 | 0 | 197 | 0 | 0 | 0 |
| 7 | 0 | 1 | 0 | 0 | 4 | 0 | 0 | 191 | 0 | 4 |
| 8 | 2 | 0 | 2 | 3 | 0 | 3 | 0 | 0 | 187 | 3 |
| 9 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 198 |

表 3: 各カテゴリごとの結果

| Category | #Data | #Correct | Accuracy |
|----------|-------|----------|----------|
| 0 | 200 | 198 | 0.990 |
| 1 | 200 | 199 | 0.995 |
| 2 | 200 | 195 | 0.975 |
| 3 | 200 | 190 | 0.950 |
| 4 | 200 | 189 | 0.945 |
| 5 | 200 | 186 | 0.930 |
| 6 | 200 | 197 | 0.985 |
| 7 | 200 | 191 | 0.955 |
| 8 | 200 | 187 | 0.935 |
| 9 | 200 | 198 | 0.990 |
| All | 2,000 | 1,930 | 0.965 |