# ROB lab 3,4 Rozpoznawanie cyfr pisanych ręcznie Paweł Paczuski (271082)

## Klasyfikacji znaków przy użyciu klasyfikatorów liniowych

Klasyfikatory liniowe zajmują się odseparowywaniem dwóch grup punktów za pomocą, w zależności od liczby wymiarów (cech), prostą, płaszczyzną, hiperpłaszczyzną. W przypadku, gdy dane zawierają więcej niż dwie grupy, np. w klasyfikacja cyfr (10 klas), stosowane są techniki mające na celu uogólnienie klasyfikacji na wiele klas przy użyciu binarnych klasyfikatorów.

Jednym z możliwych rozwiązań jest technika OvO (Onve versus One) – utworzenie $\binom{N}{2}$ klasfikatorów – wszystkich możliwych par klas. Następnie przeprowadzenie klasyfikacji danego elementu, za pomocą wszystkich uzyskanych w ten sposób $\frac{n(n-1)}{2}$ klasyfikatorów. Następnie sumowane są wyniki, zaś ostatecznym wynikiem klasyfikacji jest klasa, na którą zagłosowało najwiecej składowych klasyfikatorów. W przypadku, gdy klasyfikatory nie są zgodne, można uznać daną decyzję jako wymijającą.

Klasyfikacja znaków wiąże się z klasyfikacja danych wielowymiarowych. W przypadku danych MNIST, każdy pixel obrazka jest traktowany jako oddzielna cecha. Aby zmniejszyć wymiarowość problemu, zastosowano algorytm PCA dla 80 składowych, aby tak obrócić osie układu współrzędnych, by wzdłuż nich współrzędne miały jak największą wariancję.

## Algorytm wyznaczania parametrów płaszczyzny decyzyjnej

Płaszczyznę decyzyjną wyznaczono za pomocą metody perceptronowej opartej na bardzo prostej zależności zwiazanej z sumą elementów niepoprawnie sklasyfikowanych. Wynika to z tego, że funkcja oceny J(a) jest proporcjonalna do sumy odległości punktów źle sklasyfikowanych od powierzchni decyzyjnej. Po odpowiednich przekształceniach, metoda sformalizowana jest nastepująco:

$$a(k+1) = a(k) + \eta(k) \sum_{y \in Y} y$$

$$\eta = \frac{1}{\sqrt{k}}$$

k – nr iteracji

Y – zbiór niepoprawnie sklasyfikowanych elementów

Zastosowano zmienny współczynnik korekcji $\eta$, którego wartość zależy od nr iteracji. Gdy aktualne rozwiązanie zbliża się do lokalnego ektremum, zmniejszają

się zmiany płaszczyzny decyzyjnej, co pozwala na zbliżenie się do najlepszych wartości wag.

Optymalizacją zwiększającą dwukrotnie szybkość (dla $mincorrectionnorm = 5$) działania perceptronu (kosztem precyzji), było zatrzymanie liczenia, gdy rożnica sum błędów między iteracjami jest mniejsza niż zadana wartość.

## Jakość klasyfikacji

### Rozwiązanie referencyjne

Zbiór uczący MINST

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.9134 | 0.0572 | 0.0294 |

Zbiór testowy MINST

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.9155 | 0.0549 | 0.0296 |

### Perceptron

Przetestowano implementację perceptronu dla wszystkich par cyfr:

| plabel | nlabel | posmiss | negmiss |
|---|---|---|---|
| 0 | 1 | 0.007766 | 0.005636 |
| 0 | 2 | 0.208003 | 0.098859 |
| 0 | 3 | 0.130002 | 0.112053 |
| 0 | 4 | 0.109742 | 0.049127 |
| 0 | 5 | 0.213574 | 0.112525 |
| 0 | 6 | 0.163600 | 0.094120 |
| 0 | 7 | 0.067871 | 0.048683 |
| 0 | 8 | 0.178288 | 0.084088 |
| 0 | 9 | 0.048793 | 0.140528 |
| 1 | 2 | 0.010976 | 0.859517 |
| 1 | 3 | 0.037526 | 0.512478 |
| 1 | 4 | 0.009196 | 0.038514 |
| 1 | 5 | 0.195343 | 0.509131 |
| 1 | 6 | 0.217740 | 0.160696 |
| 1 | 7 | 0.010234 | 0.059537 |

| plabel | nlabel | posmiss | negmiss |
|---|---|---|---|
| 1 | 8 | 0.035598 | 0.780550 |
| 1 | 9 | 0.010383 | 0.036309 |
| 2 | 3 | 0.096677 | 0.089545 |
| 2 | 4 | 0.870594 | 0.303663 |
| 2 | 5 | 0.199899 | 0.340712 |
| 2 | 6 | 0.000168 | 1.000000 |
| 2 | 7 | 0.799597 | 0.296409 |
| 2 | 8 | 0.783988 | 0.393779 |
| 2 | 9 | 0.479020 | 0.317196 |
| 3 | 4 | 0.204208 | 0.205923 |
| 3 | 5 | 0.103898 | 0.845785 |
| 3 | 6 | 0.122492 | 0.074349 |
| 3 | 7 | 0.981080 | 0.020591 |
| 3 | 8 | 0.905888 | 0.117416 |
| 3 | 9 | 0.101615 | 0.622458 |
| 4 | 5 | 0.192229 | 0.553034 |
| 4 | 6 | 0.211058 | 0.131632 |
| 4 | 7 | 1.000000 | 0.000319 |
| 4 | 8 | 0.782951 | 0.085114 |
| 4 | 9 | 0.884629 | 0.083712 |
| 5 | 6 | 0.566870 | 0.108483 |
| 5 | 7 | 0.188895 | 0.139984 |
| 5 | 8 | 0.265265 | 0.192275 |
| 5 | 9 | 0.045748 | 0.793915 |
| 6 | 7 | 0.164076 | 0.348125 |
| 6 | 8 | 0.262420 | 0.600410 |
| 6 | 9 | 0.762758 | 0.460750 |
| 7 | 8 | 0.335355 | 0.220475 |
| 7 | 9 | 0.326257 | 0.364935 |
| 8 | 9 | 0.256196 | 0.638259 |

Użycie opisanej implementacji perceptronu wewnątrz komitetu OvO dla zbioru testowego MNIST dało następujące rezultaty:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.8951 | 0.0759 | 0.0288 |

Macierz pomyłek:

```
5599    0   23   12    5   61   71    4   29    2  117
   0 6376   57   25    4   20    9   45  115    5   86
  45   39 5258   69   74   24  104   54   95   16  180
```

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 24 | 49 | 101 | 5340 | 3 | 211 | 30 | 62 | 84 | 43 | 184 |
| 6 | 23 | 44 | 2 | 5272 | 12 | 76 | 16 | 29 | 205 | 157 |
| 61 | 37 | 20 | 181 | 28 | 4525 | 79 | 47 | 113 | 45 | 285 |
| 47 | 25 | 83 | 1 | 21 | 80 | 5523 | 0 | 27 | 0 | 111 |
| 12 | 42 | 64 | 21 | 46 | 14 | 5 | 5749 | 17 | 135 | 160 |
| 35 | 103 | 65 | 120 | 27 | 155 | 36 | 48 | 4908 | 79 | 275 |
| 27 | 19 | 31 | 80 | 170 | 33 | 1 | 198 | 52 | 5161 | 177 |

Zbiór testowy MNIST

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.8980 | 0.0723 | 0.0297 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 932 | 0 | 2 | 2 | 0 | 13 | 10 | 2 | 1 | 0 | 18 |
| 0 | 1074 | 4 | 7 | 0 | 1 | 4 | 2 | 31 | 0 | 12 |
| 7 | 3 | 917 | 15 | 12 | 2 | 11 | 8 | 17 | 4 | 36 |
| 4 | 2 | 14 | 907 | 0 | 23 | 2 | 18 | 14 | 3 | 23 |
| 1 | 4 | 7 | 0 | 896 | 1 | 11 | 3 | 4 | 31 | 24 |
| 11 | 5 | 5 | 35 | 3 | 738 | 12 | 8 | 16 | 10 | 49 |
| 11 | 2 | 8 | 0 | 7 | 15 | 891 | 0 | 3 | 0 | 21 |
| 1 | 10 | 19 | 5 | 7 | 0 | 1 | 931 | 4 | 23 | 27 |
| 10 | 7 | 13 | 21 | 8 | 23 | 9 | 15 | 816 | 11 | 41 |
| 5 | 3 | 5 | 8 | 33 | 5 | 1 | 23 | 2 | 878 | 46 |

Analiza macierzy pomyłek sugeruje, że komitet najgorzej sobie radzi z cyfrą 5 – jest ona mylona z 3, 8 i 6, wobec tej cyfry komitet najczęściej wstrzymał się od decyzji. Problematyczna też jest cyfra 4 – mylona jest z 6 i 8.

## Próba usprawnienia klasyfikacji

Zgodnie z sugerowanym w instrukcji do laboratorium sposobem usprawnienia klasyfikacji, dokonano próby podziału cyfy na grupy, aby następnie dokonać klasyfikacji na zmniejszonej liczbie klas wewnątrz grup.

### Podział na dwie lub trzy grupy za pomocą kmeans

Grupy zwracane przez kmeans były następnie analizowane pod kątem procentowego udziału poszczególnych klas. Poszczególne cyfry przydzielono do grupy, w ktorej występowały częściej. Zastosowano algorytm kmeans w celu podziału cyfr na dwie grupy:

g0 = [0 1 4 6 7];

g1 = [2 3 5 8 9];

Wyniki klasyfikatorów przydzielających cyfry do jednej z dwóch grup:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.9089333 | 0.0855000 | 0.0055667 |

Sumaryczne wyniki dla klasyfikatorów działających wewnątrz grup, po scaleniu wynikow:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.866100 | 0.122350 | 0.011550 |

### Podział na trzy grupy za pomocą kmeans

Podział na dwie grupy wg kmeans nie przyniósł poprawy (wręcz odwrotnie), dlatego podjęto próbę podziału cyfr na trzy grupy.

Analogicznie do poprzednego podziału zbioru cyft, przydzielono każdą z cyfr do jednej z trzech grup:

g0 = [0 2 3 5 6];

g1 = [1 8];

g2 = [4 7 9];

Wyniki klasyfikatorów przydzielających cyfry do jednej z trzech grup:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.9326333 | 0.0625333 | 0.0048333 |

Sumaryczne wyniki dla klasyfikatorów działających wewnątrz grup, po scaleniu wynikow:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.884067 | 0.103667 | 0.012267 |

### Najlepszy znaleziony podział

Aby zmaksymalizować skuteczność wstępnego podziału zbioru cyfr – znaleźć najlepszy podział, można przetestować każdy możliwy podział. Liczność możliwych podziałów zbioru opisują liczby Bella. Dla zbioru 10-elementowego (0-9) istnieje $B_{10} = 115975$ podziałów. Sprawdzenie każdego z podziałów wymaga wytrenowania klasyfikatorów na zbiorze MNIST, co zajmuje kilka

sekund. Ostatecznie, sprawdzenie wszystkich możliwości uznano za nieopłacalne czasowo. Pomimo nieopłacalności sprawdzenia wszystkich możliwości, za pomocą skryptu `autopartition`, sprawdzajacego kolejno wszystkie możliwe podziały, zwracającego na bieżąco najlepszy do tej pory znaleziony, wykryto najlepszy z testowanych podziałów:

g0 = [0 1 2 3 5 8];

g1 = [4 6];

g2 = [7 9];

Wyniki klasyfikatorów przydzielających cyfry do jednej z trzech grup:

Macierz pomyłek:

```
35046     558     378      44
  640   10766     244     110
  626     212   11321      55
```

Ocena klasyfikacji:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.9522167 | 0.0443000 | 0.0034833 |

**Szczegółowe wyniki klasyfikatorow działających wewnątrz grup**

Pierwsza grupa:

```
5695     2    32    12     0    52     0     0    25     0    34
   0  6451    41    24     0    13     0     0    90     0    57
  29    29  5353    66     0    31     0     0    72     0    96
   6    31    83  5480     0   175     0     0    73     0   115
  23    42    63    13     0     2     0     0    41     0    30
  39    14    43   161     0  4694     0     0   107     0   132
  40    35   145     2     0   126     0     0    36     0    42
  29    67    96    27     0     4     0     0    20     0    20
  17    77    69   123     0   133     0     0  5138     0   132
  42    37    51    92     0    35     0     0    75     0    31
```

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.903586 | 0.077440 | 0.018974 |

Druga grupa

```
   0    0    0    0   15    0   47    0    0    0    0
   0    0    0    0    6    0    4    0    0    0    0
   0    0    0    0  110    0   98    0    0    0    0
```

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 1 | 0 | 28 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 5243 | 0 | 33 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 57 | 0 | 106 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 44 | 0 | 5446 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 62 | 0 | 3 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 35 | 0 | 51 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 145 | 0 | 2 | 0 | 0 | 0 | 0 |

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.92658 | 0.07342 | 0.00000 |

Trzecia grupa:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 43 | 0 | 5 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 44 | 0 | 11 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 68 | 0 | 71 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 35 | 0 | 209 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0 | 49 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5739 | 0 | 191 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 0 | 52 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 209 | 0 | 5182 | 0 |

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.91443 | 0.08557 | 0.00000 |

Elementy zbioru, których przynależności do grup wstępny przydział nie był w stanie rozstrzygnąć, zostały sklasyfikowane za pomocą pełnego zbioru klasyfikatorów:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| 0 | 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 0 | 3 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 6 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 96 | 0 | 0 | 1 | 0 | 3 | 7 |
| 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 3 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 1 | 2 | 0 | 1 | 1 | 0 | 3 |
| 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 1 | 2 |
| 0 | 0 | 0 | 0 | 16 | 2 | 0 | 0 | 1 | 21 | 7 |

Sumaryczne wyniki dla klasyfikatorów działających wewnątrz grup, po scaleniu wynikow:

| OK | Błąd | Odrzucenie |
|---|---|---|
| 0.910187 | 0.078289 | 0.011523 |

Sumaryczne wyniki są nieco lepsze w stosunku do rozwiązania z użyciem komitetu z własną implementacją perceptronu, jednak nie można stwierdzić, że są one lepsze w stosunku do rozwiązania referencyjnego (przyczyną może być np. optymalizacja działania perceptronu polegająca na zakończeniu obliczania, gdy zmiany płaszczyzny są mniejsze niż określona ręcznie wartość, która przyspiesza ok. dwukrotnie działanie programu).

Skuteczność tego rozwiązania oparta jest w dużej mierze na wysokiej skuteczności klasyfikatorów binarnych w grupach 2 i 3. Przez to, że te grupy zawierają tylko dwie cyfry, to są to pojedyncze kanoniczne klasyfikatory wyciągnięte poza komitet głosujący, nie mogą wycofać się z głosowania, co okazuje się, że ma pozytywny wpływ na sumaryczny wynik. Możliwe, że istnieje jeszcze lepszy sposób podziału zbioru cyfr, który pozwoliłby na uzyskanie znacznej poprawy jakości klasyfikacji.