

## WSI – zadanie 5

- Wstęp

Zadaniem było zaimplementowanie maszyny wektorów nośnych, aby klasyfikowała cyfry ze zbioru MNIST.

- Rozwiązanie

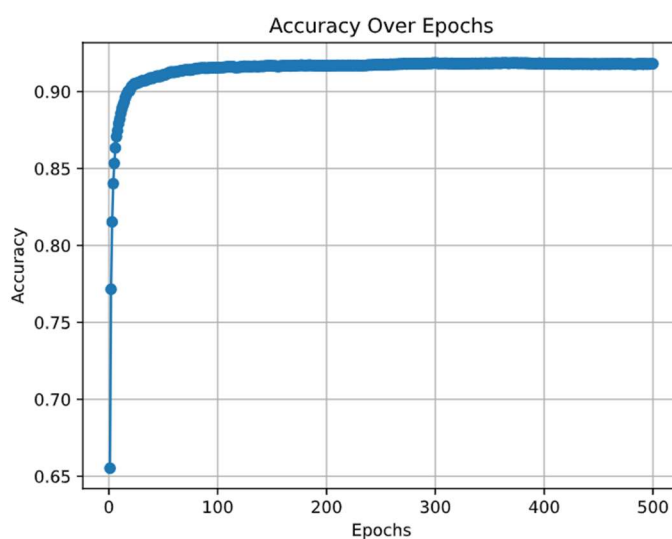
Pojedyncza maszyna wektorów nośnych jest klasyfikatorem binarnym. Żeby zakwalifikować cyfrę do jednej z 10 kategorii, to moje rozwiązanie polega na 10 maszynach. Każda z nich odpowiada jednej cyfrze.

Klasyfikacja liczb jest uzależniona od tabeli wag każdej maszyny, oraz jej biasu. Po zsumowaniu iloczynów każdego piksela i odpowiadającej jemu wagi, oraz odjęciu biasu, otrzymujemy wynik. Następnie porównujemy, która z maszyn osiągnęła największy wynik, no i której cyfrze ta maszyna odpowiada.

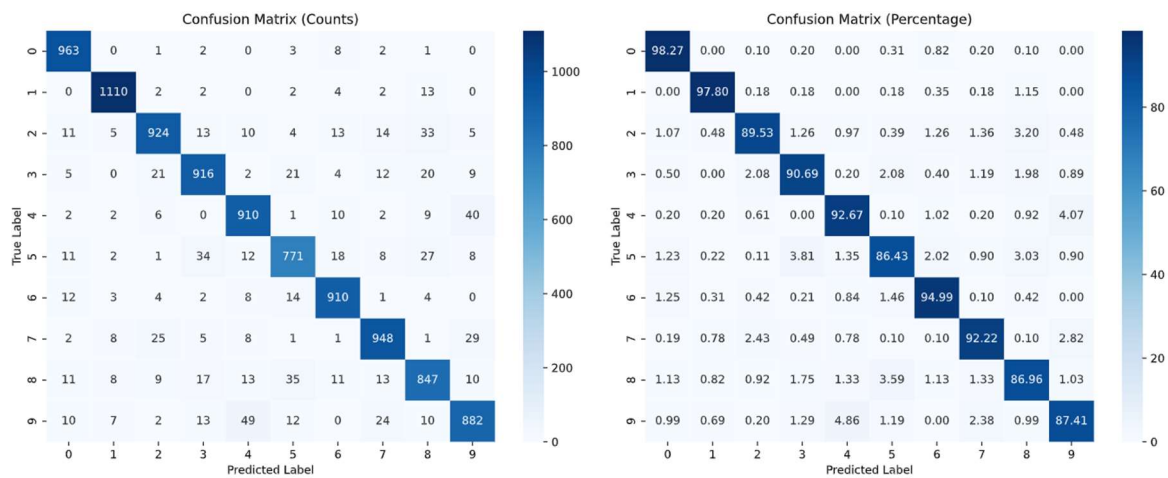
Trening polega na takim dobraniu wag oraz biasu dla każdej z maszyn, aby po podaniu obrazka maszyna prawidłowo klasyfikowała go jako cyfrę. Osiąga się to poprzez przepuszczanie przez nią obrazków treningowych, wraz z ich identyfikatorami. Najpierw liczymy poprzednio wspomniany wynik dla danego obrazka. Następnie mnoży się go przez 1 lub -1, w zależności od tego, czy obrazek treningowy jest lub nie jest cyfrą, którą maszyna ma rozpoznawać. Jeżeli wynik jest mniejszy od 1, to znaczy, że maszyna błędnie zaklasyfikowała obrazek, więc należy zmodyfikować wagi oraz bias. To, jak mocno jeden obrazek wpływa na ich modyfikację, można regulować parametrami `learning_rate`, oraz `lambda`. W moim rozwiązaniu jedna epoka polega na wykonaniu takiego treningu używając całego zbioru obrazków treningowych dla każdej z maszyn.

- Wyniki

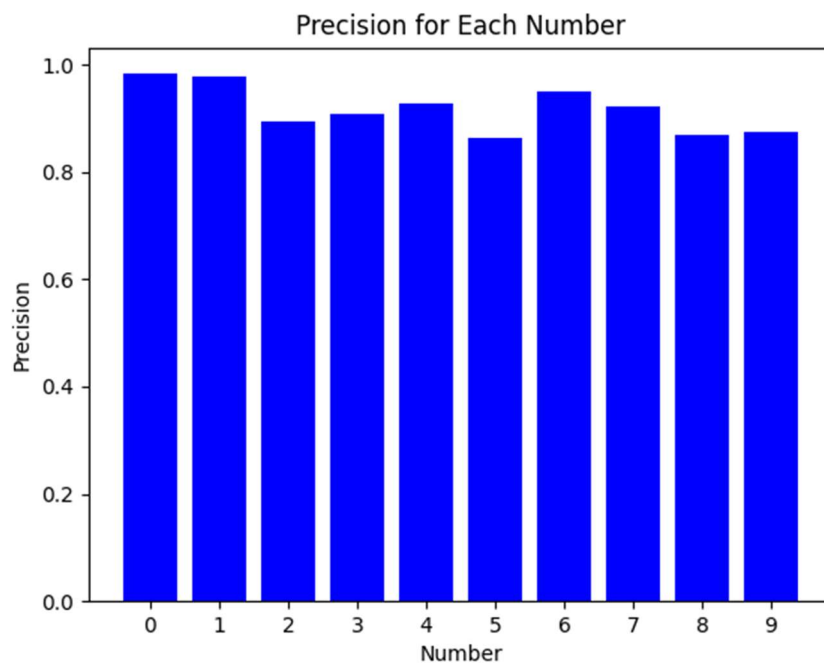
Najlepsze wyniki udawało mi się osiągać dla parametrów: ilość epok – 500,  $\lambda = 0.01$ , `learning_rate` = 0.0001:

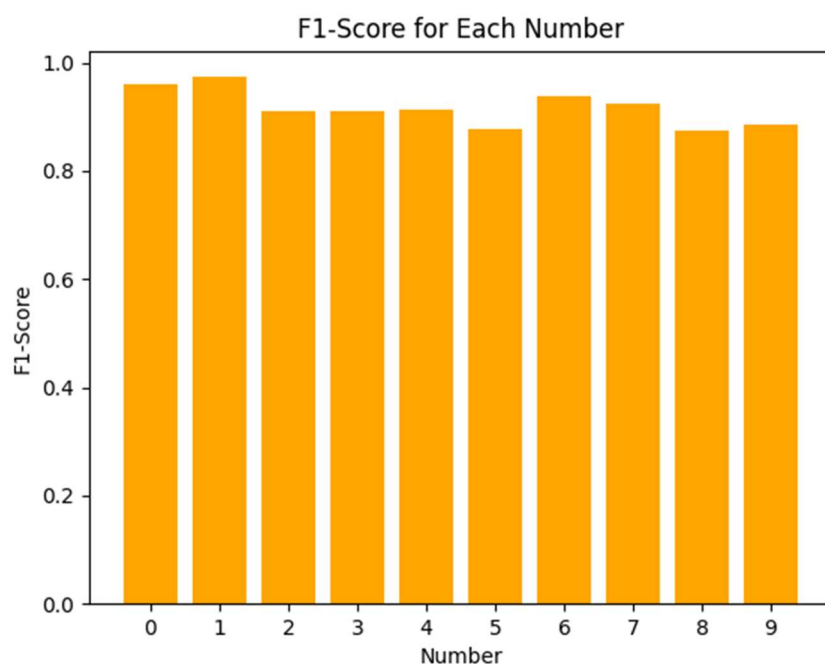
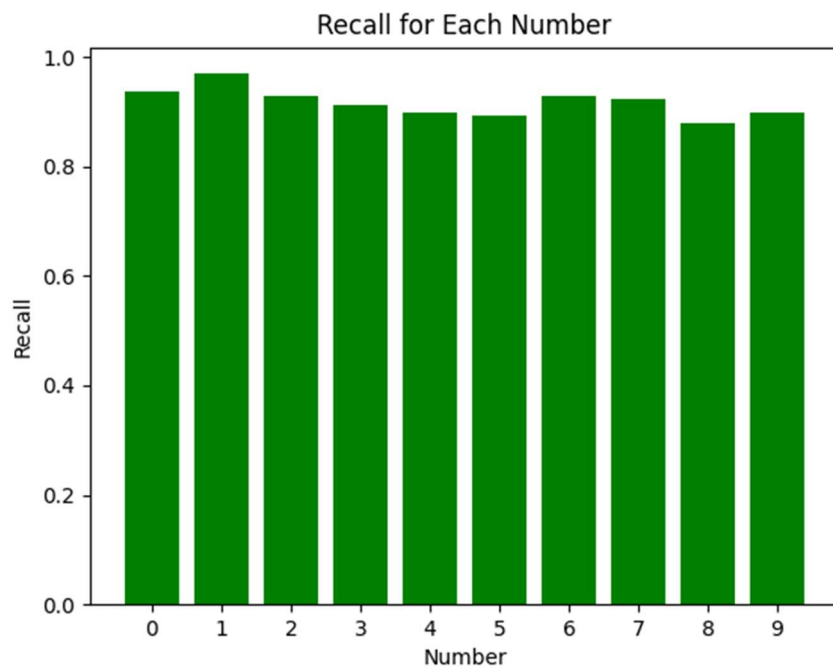


Kształt wykresu jest zbliżony do logarytmicznego. Po pierwszych około 20 epokach maszyna osiągnęła skuteczność 90%, a ostatecznie skończyła na 91,8%.



Tak wygląda confusion matrix. Z niego widać, że maszyna osiąga największą skuteczność dla cyfry 0, a najgorszą dla 5.





Tak wyglądają wykresy słupkowe dla precyzji, rozrzutów i f1-score dla każdej z cyfr. Co ciekawe, najwyższy f1-score ma cyfra 1, a nie 0, za to najniższy – 8, a nie 5. Różnice są minimalne, ale pokazują, że branie pod uwagę tylko proporcji prawidłowych klasyfikacji do wszystkich klasyfikacji nie zawsze dają pełny obraz. Przykładowo – gdy maszyna otrzymywała obrazek 0, to w 98,3% przypadków klasyfikowała go jako cyfrę 0. Cyfra 1 była prawidłowo klasyfikowana podobnie, bo w 97,8% przypadków. Natomiast, maszyna była częściej skłonna błędnie przypisać cyfrę 0 innej cyfrze, niż cyfrę 1. Dlatego rozrzut cyfry 1 jest o ok. 0,03 większy, co składa się na wyższy f1-score dla 1 o ok. 0,02.