**Sprawozdanie –Techniki Optymalizacji**

Autorzy:  
Paweł Rosolak 100810  
Tomasz Straszewski 101319  
Grupa TWO

1. **Algorytm regresji logistycznej zaimplementowany metodą Newtona-Raphsona.**
   1. Liczba iteracji: 8, czas obliczeń: 37.5s
   2. Obliczone błędy:

|  |  |  |
| --- | --- | --- |
| **Rodzaj błędu** | **Zbiór** | **Wartość** |
| Zero-jedynkowy | Treningowy | 0.1346 |
| Zero-jedynkowy | Testowy | 0.1682 |
| Logistyczny | Treningowy | 0.3063 |
| Logistyczny | Testowy | 0.3796 |

* 1. Histogram wartości wektora wag:

1. **Algorytm regresji logistycznej metodą stochastycznego spadku wzdłuż gradientu.**
   1. Długość kroku obliczana jest z następującego wzoru:

Gdzie:  
**x** – Liczba obserwacji

**i** – aktualna iteracja w danej epoce

Długość kroku wyznaczana została zgodnie z powyższym wzorem, ponieważ tylko w jego przypadku udało nam się przejść do kolejnych epok i otrzymać niższe wartości błędów. Wzór został wyznaczony eksperymentalnie.

* 1. Liczba epok: 13, czas obliczeń, ok. 18min.
  2. Obliczone błędy:

|  |  |  |
| --- | --- | --- |
| **Rodzaj błędu** | **Zbiór** | **Wartość** |
| Zero-jedynkowy | Treningowy | 0.1941 |
| Zero-jedynkowy | Testowy | 0.2169 |
| Logistyczny | Treningowy | 2.3467 |
| Logistyczny | Testowy | 2.9675 |

* 1. Wykresy:
  2. Histogram wartości wektora wag:

1. **Wnioski.**
   1. Dla tego zestawu danych lepsza wydaje się metoda Newtona-Raphsona. Jest ona szybsza oraz daje lepsze wyniki w stosunku do metody stochastycznego spadku wzdłuż gradientu.
   2. W obu metodach błąd liczony na zbiorze treningowym zawsze jest mniejszy. Ten fakt nie jest zaskakujący. Model został utworzony na zbiorze treningowym, a zatem intuicja sama nam podpowiada, że błąd liczony na tym zbiorze będzie mniejszy niż na zbiorze testowym.
   3. W obu przypadkach nie udało się nam zauważyć zbieżności błędu w funkcji liczby epok. Oba błędy nie są uzależnione od ilości epok przez co można przypuszczać, że w poszczególnych epokach ma miejsce „przeuczenie” modelu.
   4. W obu przypadkach rozkład wag jest bardzo podobny. Wartości skupiają się w dwóch środkowych przedziałach.