

SIECI NEURONOWE

Sprawozdanie - Ćwiczenie 1

Aleksander Poławski
Grupa - Poniedziałek 18:15
Prowadzący - mgr inż. Jan Jakubik

18 października, 2020

1 Cel ćwiczenia

Celem ćwiczenia pierwszego laboratoriów kursu Sieci Neuronowe było poznanie podstawowych funkcji wykonywanych przez pojedynczy neuron, obserwacja zachowania neuronu przy różnych funkcjach przejścia oraz określenie wielkości, które mają wpływ na szybkość uczenia neuronu.

2 Plan ćwiczenia oraz badań

- a) stworzenie programu symulującego działanie pojedynczego neuronu, w formie perceptronu prostego realizującego logiczną funkcję AND oraz przeprowadzenie eksperymentów badających szybkość i skuteczność uczenia się tego perceptronu w zależności od:
 - wartości progu aktywacji i użycia dynamicznego progu (bias)
 - zakresu wartości początkowych losowych wag
 - wartości współczynnika uczenia α
 - zastosowanej funkcji przejścia neuronu (funkcji progowej unipolarnej lub funkcji progowej bipolarnej)
- b) modyfikacja zaimplementowanego perceptronu prostego do Adaline oraz przeprowadzenie podobnych badań jak w przypadku perceptronu prostego - tj. badań szybkości uczenia się Adaline w zależności od zakresu początkowych, losowych wag oraz współczynnika uczenia α
- c) porównanie własności, skuteczności, wyników badań perceptronu prostego i Adaline

3 Opis zaimplementowanego programu

Program zaimplementowano w środowisku Visual Studio w języku C#. Składa się z następujących klas:

- DataSets - umożliwia przechowywanie zbiorów uczących i testowych. Klasa pozwala na definiowanie i obsługę zbiorów o dowolnej liczbie wektorów wejściowych o dowolnej długości wektora wejściowego
- Entry - klasa opisująca pojedynczy wektor wejściowy
- PerceptronSettings - instancje tej klasy pozwalają na przechowywanie ustawień perceptronu takich jak: zakres wartości wag początkowych, rodzaj funkcji przejścia, użycie dynamicznego biasu, wartość progu aktywacji, rodzaj perceptronu (prosty lub adaline), wartość współczynnika uczenia α itd.
- Perceptron - klasa zawierająca logikę uczenia perceptronu o ustawieniach zdefiniowanych w instancji PerceptronSettings
- Program - klasa główna - manager - organizuje kolejność wykonywania zadań programu, zawiera predefiniowane testy potrzebne do wykonania badań przewidzianych w ćwiczeniu

Program w sposób prosty i intuicyjny umożliwia wykonanie wszystkich zaplanowanych w ćwiczeniu badań. Jego zaletą jest też możliwość rozwiązywania innych problemów niż "AND", dzięki implementacji obsługi danych wejściowych o różnych długościach wektorów wejściowych.

4 Badania

W poniższej sekcji zamieszczono i opisano wyniki badań.

4.1 Wpływ wartości progu aktywacji na wydajność uczenia perceptronu prostego

a) Założenia

- rodzaj perceptronu: [prosty]
- funkcja przejścia: [unipolarna]
- zakres wartości początkowych wag: od [-0.8] do [0.8]
- wartość współczynnika uczenia α : [0.01]
- brak dynamicznego progu - biasu
- **zmienna wartość progu aktywacji: [-0.50], [-0.25], [0.00], [0.25], [0.50], [0.75], [0.90]**
- koniec uczenia po 10000 iteracjach lub przy sumie błędów równej 0 dla całej epoki

b) Przebieg

- Dla każdej wartości progu wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.
- Dodatkowo, obok badań dla danych wejściowych wyrażenia "AND" wykonano także badania dla wyrażenia "OR" w celu potwierdzenia czy optymalna wartość progu zależy od zestawu wejściowego.

c) Otrzymane wyniki

Próg aktywacji	AND - Średnia liczba iteracji ze 100 prób	OR - Średnia liczba iteracji ze 100 prób
-0.50	nie wyuczono	nie wyuczono
-0.25	nie wyuczono	nie wyuczono
0.00	nie wyuczono	nie wyuczono
0.25	65.45	54.82
0.50	55.98	67.56
0.75	57.54	92.45
0.90	46.02	103.20

Tabela 1: wpływ wartości progu aktywacji na wydajność uczenia perceptronu prostego

d) Komentarz

- Badania pokazały, że wydajność perceptronu prostego jest silnie zależna od dobranej wartości progu aktywacji. Ponadto, optymalna wartość tego progu jest również mocno zależna od charakterystyki zestawu danych treningowych.
- Źle dobrany próg jest w stanie całkowicie uniemożliwić wyuczenie perceptronu.
- Wykonano podobne badanie tym razem używając dynamicznego progu (bias). Pomyślne wyuczenie następuje w tej konfiguracji średnio po 60 iteracjach (dla "OR" i "AND").
- Zalety wykorzystania automatycznego progu są nieocenione, a spadek wydajności znikomy. Alternatywą jego użycia byłoby strojenie progu aktywacji każdorazowo przy zmianie charakterystyki danych treningowych.

4.2 Wpływ wartości zakresu wag początkowych na wydajność uczenia perceptronu prostego

a) Założenia

- rodzaj perceptronu: [prosty]
- funkcja przejścia: [unipolarna]
- wartość współczynnika uczenia α : [0.01]
- dynamiczny próg - bias
- **zmienny zakres wartości początkowych wag: [-1.0] do [1.0], [-0.8] do [0.8], [-0.6] do [0.6], [-0.4] do [0.4], [-0.2] do [0.2], [-0.1] do [0.1], [-0.01] do [0.01]**
- koniec uczenia po 10000 iteracjach lub przy sumie błędów równej 0 dla całej epoki

b) Przebieg

- Dla każdej wartości zmiennej wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.

c) Otrzymane wyniki

Zakres wag początkowych	Średnia liczba iteracji ze 100 prób
-1.0 do 1.0	67.88
-0.8 do 0.8	54.08
-0.6 do 0.6	36.66
-0.4 do 0.4	28.96
-0.2 do 0.2	15.79
-0.1 do 0.1	9.09
-0.01 do 0.01	3.88

Tabela 2: wpływ wartości zakresu wag początkowych na wydajność uczenia perceptronu prostego

d) Komentarz

- Eksperyment wykazał, że najwydajniejsze w tym wypadku jest ustawienie jak najmniejszej wartości wag początkowych
- Sprawdzono dalej, że tendencja ta nie zależy od rodzaju danych wejściowych ("AND"/"OR")
- Dalsze badania wykazały, że istotna jest korelacja wag ze współczynnikiem α , najlepsze wyniki osiągane są, jeśli te dwa parametry są podobnego rzędu wielkości, co tłumaczy otrzymane wyniki badania

4.3 Wpływ wartości współczynnika uczenia α na wydajność uczenia perceptronu prostego

a) Założenia

- rodzaj perceptronu: [prosty]
- funkcja przejścia: [unipolarna]
- **wartość współczynnika uczenia α : [0.01], [0.05], [0.2], [0.5], [0.9], [1.00], [10.00]**
- dynamiczny próg - bias
- zakres wartości początkowych wag: od [-0.5] do [0.5]
- koniec uczenia po 10000 iteracjach lub przy sumie błędów równej 0 dla całej epoki

b) Przebieg

- Dla każdej wartości zmiennej wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.

c) Otrzymane wyniki

Wartość współczynnika α	Średnia liczba iteracji ze 100 prób
0.01	30.84
0.05	10.44
0.20	6.09
0.50	6.31
0.90	5.92
1.00	5.90

Tabela 3: wpływ wartości współczynnika uczenia α na wydajność uczenia perceptronu prostego

d) Komentarz

- Eksperyment potwierdza korelację parametru z zakresem wag początkowych stwierdzoną we wnioskach badania poprzedniego (istotne jest wzajemne dobranie wag początkowych i skoku uczenia)
- Zupełnie inne wnioski wynikają z badań przy wyłączonym automatycznym progu (biasie). Okazuje się, że oprócz korelacji z początkowym zakresem wag istnieje silna korelacja z dobieranym progiem aktywacji. Jest to kolejne potwierdzenie ogromnej roli biasu w implementacji perceptronu.

4.4 Wpływ funkcji przejścia na wydajność uczenia perceptronu prostego

a) Założenia

- rodzaj perceptronu: [prosty]
- **zmienna funkcja przejścia: [unipolarna], [bipolarna]**
- wartość współczynnika uczenia α : [0.01]
- dynamiczny próg - bias
- zakres wartości początkowych wag: od [-0.5] do [0.5]
- koniec uczenia po 10000 iteracjach lub przy sumie błędów równej 0 dla całej epoki

b) Przebieg

- Dla każdej wartości zmiennej wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.

c) Otrzymane wyniki

Funkcja przejścia	AND - Średnia liczba iteracji ze 100 prób	OR - Średnia liczba iteracji ze 100 prób
unipolarna	37.34	37.20
bipolarna	9.15	9.34

Tabela 4: wpływ funkcji przejścia na wydajność uczenia perceptronu prostego

d) Komentarz

- Eksperyment wykazał, że przy podobnych ustawieniach funkcja przejścia bipolarna jest efektywniejsza od unipolarnej niezależnie od postawionego problemu
- Badania powtórzone przy innych ustawieniach wag oraz współczynnika uczenia i tendencja potwierdziła się

4.5 Wpływ wartości zakresu wag początkowych na wydajność uczenia perceptronu Adaline

a) Założenia

- rodzaj perceptronu: [Adaline]
- funkcja przejścia: [bipolarna]
- wartość współczynnika uczenia α : [0.01]
- dynamiczny próg - bias
- **zmienny zakres wartości początkowych wag:** [-1.0] do [1.0], [-0.8] do [0.8], [-0.6] do [0.6], [-0.4] do [0.4], [-0.2] do [0.2], [-0.1] do [0.1], [-0.01] do [0.01]
- koniec uczenia po 10000 iteracjach lub przy LMS równym 0.3 dla całej epoki

b) Przebieg

- Dla każdej wartości zmiennej wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.

c) Otrzymane wyniki

Zakres wag początkowych	Średnia liczba iteracji ze 100 prób
-1.0 do 1.0	42.59
-0.8 do 0.8	40.27
-0.6 do 0.6	37.5
-0.4 do 0.4	36.65
-0.2 do 0.2	36.27
-0.1 do 0.1	36.38
-0.01 do 0.01	36.02

Tabela 5: wpływ wartości zakresu wag początkowych na wydajność uczenia perceptronu Adaline

d) Komentarz

- Badania wykazały podobne zależności jak w przypadku analogicznych badań dla perceptronu prostego (sekcja 4.2)
- Wpływ zakresu wag początkowych wydaje się w tym wypadku (adaline) mniejszy, jednak fakt ten jest prawdopodobnie silnie uwarunkowany inną funkcją wyliczania błędu (rzeczywiste wartości - mniejsza korekcja wag w kroku)

4.6 Wpływ wartości współczynnika uczenia α na wydajność uczenia perceptronu Adaline

a) Założenia

- rodzaj perceptronu: [Adaline]
- funkcja przejścia: [bipolarna]
- **wartość współczynnika uczenia α :** [0.01], [0.02], [0.03], [0.05], [0.055], [0.060], [0.07]
- dynamiczny próg - bias
- zakres wartości początkowych wag: od [-0.5] do [0.5]
- koniec uczenia po 10000 iteracjach lub przy LMS równym 0.3 dla całej epoki

b) Przebieg

- Dla każdej wartości zmiennej wykonanych zostało 100 procedur uczenia, a następnie (dla każdej wartości) obliczona została średnia ilość iteracji potrzebnych do wyuczenia perceptronu.

c) Otrzymane wyniki

Wartość współczynnika α	Średnia liczba iteracji ze 100 prób
0.010	35.15
0.020	21.10
0.030	14.07
0.050	11.31
0.055	11.12
0.060	nie wyuczono

Tabela 6: wpływ wartości współczynnika uczenia α na wydajność uczenia perceptronu Adaline

d) Komentarz

- Badania wykazały, że Adaline jest silnie zależny od współczynnika uczenia mimo użycia dynamicznego progu (bias)
- Zbyt duży współczynnik powoduje za duży krok uczenia i ciągle (cykliczne) 'przeskakiwanie' optymalnych wartości wag, a w wyniku nie osiągnięcie wyznaczonego progu błędu średniokwadratowego i pożądanego wyników
- Z drugiej strony, im większy współczynnik tym większa szansa na szybsze znalezienie optimum

4.7 Wpływ wartości dopuszczalnego błędu na wydajność i skuteczność uczenia perceptronu Adaline

Badania wykazały, że zgodnie z intuicją proces nauczania Adaline jest krótszy dla większych wartości dopuszczalnego błędu.

W przeciwieństwie do perceptronu prostego nie jest jednak możliwe stwierdzenie, że Adaline został skutecznie wyuczony (przed dokonaniem testów). Im większy dopuszczalny błąd średniokwadratowy tym większa szansa, że uczenie zakończy się niepowodzeniem.

4.8 Porównanie wyników badań Adaline i perceptronu prostego

Szczegółowe porównanie zostało przedstawione w poszczególnych sekcjach sprawozdania. Poniżej przedstawiono wybrane różnice i podobieństwa oraz wnioski wynikające z dodatkowych badań:

- Użycie Adaline nie daje pewności wyuczenia dla określonych danych treningowych (w przeciwieństwie do perceptronu prostego). Z drugiej strony, jeśli nie uda się dobrze dostroić perceptronu prostego, nie otrzymamy prawie żadnych wyników, a w wypadku Adaline mimo braku całkowitego wyuczenia otrzymamy działający (w jakimś stopniu) model.
- Niezależnie od użytego perceptronu zmiana zakresu wag początkowych, współczynnika uczenia lub progu aktywacji powoduje podobne zmiany wydajności (ważniejsza jest wzajemna zależność między wymienionymi parametrami).
- Nie stwierdzono wyższości jednego z perceptronów nad drugim w kontekście wydajności. Zależy ona głównie od innych parametrów.
- W obu przypadkach użycie automatycznego progu (biasu) wiąże się z ogromnymi korzyściami i ułatwia strojenie innych parametrów.
- Po wykonaniu dodatkowych badań, korzystając z danych treningowych opartych na "XOR", potwierdzony został fakt, że Adaline i perceptron prosty są w stanie rozwiązać tylko problemy separowalnie liniowe.

5 Podsumowanie

Pomyślnie udało się zrealizować wytyczne zadania i dokonać ciekawych badań. W wyniku poznano elementarne pojęcia dotyczące sieci neuronowych oraz nauczono się mechanizmów działania perceptronu prostego i perceptronu Adaline.