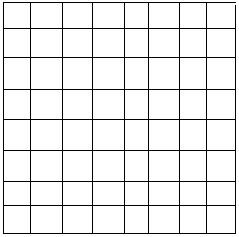
**Rafał Szyszka**

**Sprawozdanie 4 – Uczenie sieci regułą Hebba.**

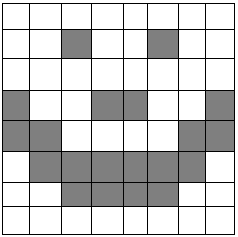
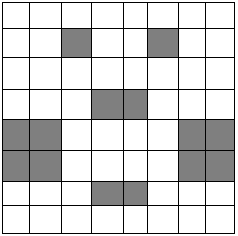
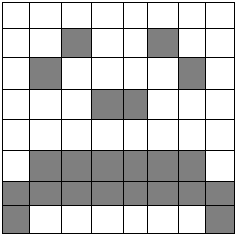
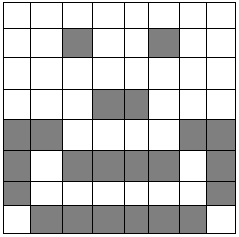
1. **Cel ćwiczenia:**

Celem ćwiczenia jest poznanie działania reguły Hebba na przykładzie rozpoznawania emotikon.

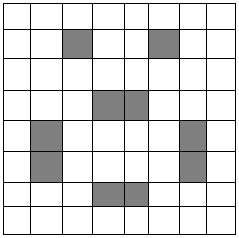
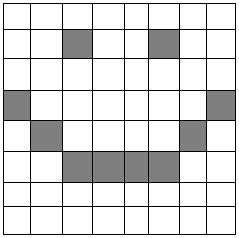
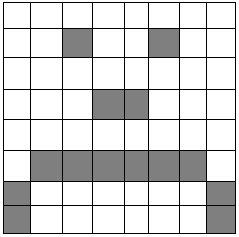
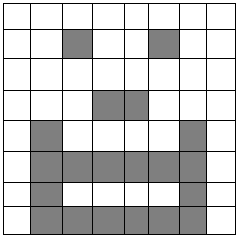
1. **Przebieg wykonania ćwiczenia**
2. Przygotowanie danych uczących i testujących dla 4 różnych emotikonek na siatce o wymiarze 8x8



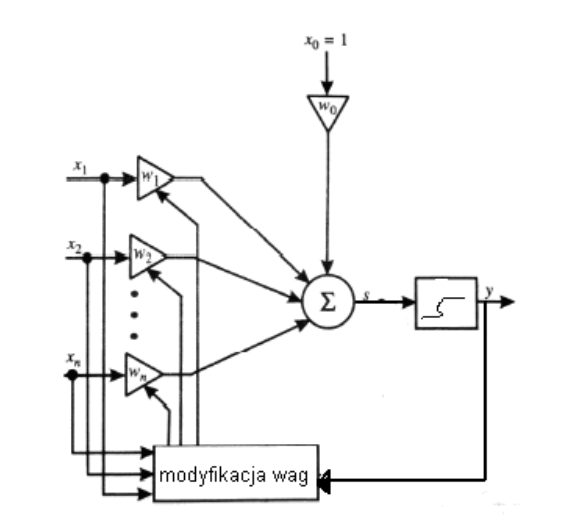
**Dane uczące:**



**Dane testujące:**



1. Implementacja sieci oraz reguły Hebba w wersji z i bez współczynnika zapominania
2. Uczenie sieci dla różnych współczynników uczenia i zapominania
3. Testowanie sieci
4. **Syntetyczny opis budowy użytej sieci i algorytmów uczenia:**



*Rys. 1 – Model neuronu Hebba*

Ogólny model neuronu Hebba, przedstawiony na rys1. odpowiada standardowej postaci modelu neuronu. Waga w­­ij włączona jest między sygnałem wejściowym yj a węzłem sumacyjnym i-tego neuronu o sygnale wyjściowym yj. Uczenie neuronu z zastosowaniem reguły Hebba w trybie bez nauczyciela polega na używaniu aktualnej wartości yi sygnału wyjściowego neuronu.

Jeśli j-ta komórka o sygnale wyjściowym yj powiązana jest z i-tą o sygnale wyjściowym yi przez wagę wij to na stan powiązań tych komórek wpływają wartości sygnałów wyjściowych yj oraz yi.Zmiana wagi wij odbywa się proporcjonalnie do iloczynu jego sygnału wejściowego oraz wyjściowego.

gdzie η jest stałą uczenia z przedziału 0 – 1. yj – sygnał wyjściowy neuronu yi – wartość wejściowa.

Reguła Hebba charakteryzuje się tym, że w jej wyniku wagi mogą przybierać wartości dowolnie duże, gdyż w każdym cyklu uczącym następuje proces sumowania aktualnych wartości wag i skończonego przyrostu Δwij

Jedną z metod poprawy stabilności procesu uczenia wg reguły Hebba jest przyjęcie przy aktualizacji wag nie ostatnie wartości wij, ale wartości zmniejszonej o tak zwany współczynnik zapominania γ. Wówczas regułę Hebba można zapisać w postaci:

Współczynnik zapominania γ zawiera się zwykle w przedziale 0 – 1.

1. **Zestawienie i analiza otrzymanych wyników.**
2. Wyniki dla przypadku z współczynnikiem zapominania

*Tabela 1 – zbiorcze zestawienie wyników.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | lr | 0,01 |  | lr | 0,05 |  | lr | 0,005 |  | lr | 0,0075 |  | lr | 0,025 |
|  | fr | 0,003 |  | fr | 0,016 |  | fr | 0,0016 |  | fr | 0,0025 |  | fr | 0,0083 |
| nr testu | l. epok | l. popr. |  | l. epok | l. popr. |  | l. epok | l. popr. |  | l. epok | l. popr. |  | l. epok | l. popr. |
| 1 | 193 | 75% |  | 34 | 25% |  | 10 | 50% |  | 9 | 75% |  | 8 | 50% |
| 2 | 206 | 25% |  | 1 | 50% |  | 12 | 0% |  | 276 | 50% |  | 90 | 75% |
| 3 | 1 | 50% |  | 6 | 25% |  | 19 | 25% |  | 17 | 0% |  | 1 | 25% |
| 4 | 10 | 50% |  | 3 | 75% |  | 13 | 25% |  | 288 | 50% |  | 5 | 50% |
| 5 | 10 | 75% |  | 5 | 50% |  | 434 | 50% |  | 225 | 25% |  | 4 | 0% |
| 6 | 14 | 25% |  | 37 | 50% |  | 162 | 25% |  | 142 | 25% |  | 4 | 50% |
| 7 | 208 | 75% |  | 1 | 0% |  | 322 | 25% |  | 12 | 75% |  | 5 | 25% |
| 8 | 12 | 100% |  | 1 | 25% |  | 178 | 0% |  | 345 | 25% |  | 38 | 50% |
| 9 | 17 | 25% |  | 2 | 50% |  | 249 | 75% |  | 275 | 50% |  | 6 | 25% |
| 10 | 212 | 50% |  | 2 | 0% |  | 414 | 25% |  | 23 | 50% |  | 27 | 25% |
| **ŚREDNIO** | **88,3** | **55%** |  | **9,2** | **35%** |  | **181,3** | **30%** |  | **161,2** | **43%** |  | **18,8** | **38%** |

Tabela 1 przestawia zbiorcze zestawienie wyników dla różnych współczynników uczenia i zapominania. Dla każdego zestawu współczynników zostało wykonane 10 testów, w których zmierzone zostały liczba epok potrzebnych do nauczenia sieci, oraz % poprawnych odpowiedzi dla danych testujących.

|  |  |  |  |
| --- | --- | --- | --- |
| lr | fr | l. epok | %poprawnych |
| 0,005 | 0,0016 | 181,3 | 30% |
| 0,0075 | 0,0025 | 161,2 | 43% |
| 0,01 | 0,003 | 88,3 | 55% |
| 0,025 | 0,0083 | 18,8 | 38% |
| 0,05 | 0,016 | 9,2 | 35% |

*Tabela 2 – średnie wyniki dla danych współczynników uczenia i zapominania*

W opracowaniu wyników zostało przyjęte, że współczynnik zapominania stanowi współczynnika uczenia. Na powyższej tabelce widzimy, że średnia epok wraz ze wzrostem learning rate i forget rate maleje. Zaś % poprawnych wyników dla danych testujących wskazuje, że optymalny współczynnik uczenia oscyluje w granicach 0,01. Powyższe wyniki przedstawiono także na wykresie 1 i 2.

*Wykres 1.*

Na powyższym wykresie jasno widać, że liczba epok wraz ze wzrostem współczynników uczenia i zapominania maleje.

*Wykres 2.*

Tutaj zauważamy, jak znalezione zostało optymalne ustawienie współczynników uczenia i zapominania dla naszej sieci i danych uczących i testujących.

Analizując powyższe wyniki widzimy, że mimo ciągłego spadku liczby epok potrzebnych do nauczenia naszej sieci, liczba popełnianych błędów była najmniejsza dla współczynnika uczenia 0,01 i jego odpowiednika współczynnika zapominania 0,0033.

1. Wyniki dla przypadku bez współczynnika zapominania.

*Tabela 3. – wyniki zbiorcze.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | lr | 0,005 |  | lr | 0,0075 |  | lr | 0,01 |  | lr | 0,025 |  | lr | 0,05 |
| nr testu | l. epok | % popr. |  | l. epok | % popr. |  | l. epok | % porp. |  | l epok | % popr. |  | l. epok | % popr. |
| 1 | 9 | 25% |  | 19 | 50% |  | 20 | 50% |  | 3 | 50% |  | 3 | 25% |
| 2 | 434 | 0% |  | 12 | 75% |  | 213 | 75% |  | 85 | 50% |  | 45 | 25% |
| 3 | 49 | 25% |  | 13 | 50% |  | 13 | 50% |  | 1 | 50% |  | 2 | 50% |
| 4 | 17 | 75% |  | 280 | 50% |  | 8 | 50% |  | 3 | 75% |  | 3 | 50% |
| 5 | 29 | 25% |  | 6 | 0% |  | 145 | 25% |  | 5 | 25% |  | 6 | 25% |
| 6 | 256 | 100% |  | 19 | 25% |  | 10 | 25% |  | 11 | 0% |  | 45 | 50% |
| 7 | 9 | 50% |  | 270 | 75% |  | 216 | 25% |  | 88 | 50% |  | 2 | 25% |
| 8 | 428 | 75% |  | 98 | 50% |  | 202 | 25% |  | 6 | 25% |  | 3 | 50% |
| 9 | 1 | 50% |  | 294 | 50% |  | 12 | 0% |  | 4 | 25% |  | 46 | 75% |
| 10 | 48 | 0% |  | 2 | 25% |  | 1 | 25% |  | 92 | 25% |  | 2 | 50% |
| **SREDNIO** | **128** | **43%** |  | **101,3** | **45%** |  | **84** | **35%** |  | **29,8** | **38%** |  | **15,7** | **43%** |

*Tabela 4 – średnie wyniki dla danego współczynnika uczenia*

|  |  |  |
| --- | --- | --- |
| lr | l. epok | %poprawnych |
| 0,005 | 128 | 43% |
| 0,0075 | 101,3 | 45% |
| 0,01 | 84 | 35% |
| 0,025 | 29,8 | 38% |
| 0,05 | 15,7 | 43% |

Na podstawie powyższej tabeli zdecydowanie możemy stwierdzić, że wysokość poprawnych wyników sieci jest niższa niż dla wariantu z współczynnikiem zapominania. W tym przypadku % poprawnych wyników nie przekroczył 50%.

*Wykres 3.*

Tutaj podobnie jak w wariancie z współczynnikiem zapominania liczba epok potrzebnych do nauczenia sieci ciągle spada.

*Wykres 4.*

Na powyższym wykresie widzimy, że % poprawnych wyników w wariancie bez współczynnika zapominania dla danych testujących nie przekracza 50%, co wyraźnie sugeruje, że wariant z współczynnikiem zapominania pozwala osiągnąc lepsze wyniki.

1. **Wnioski.**

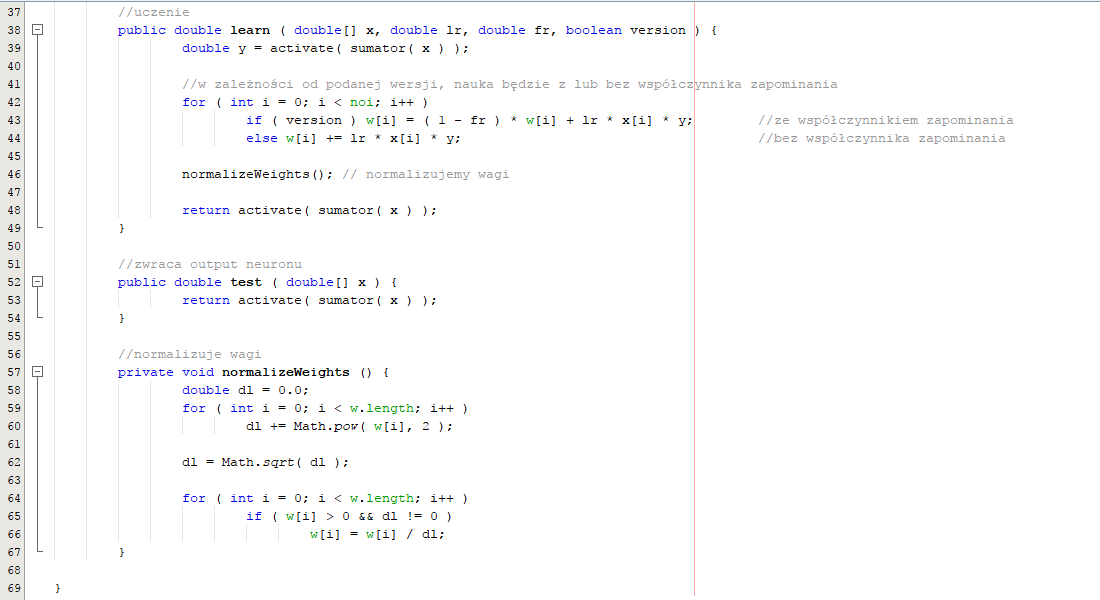
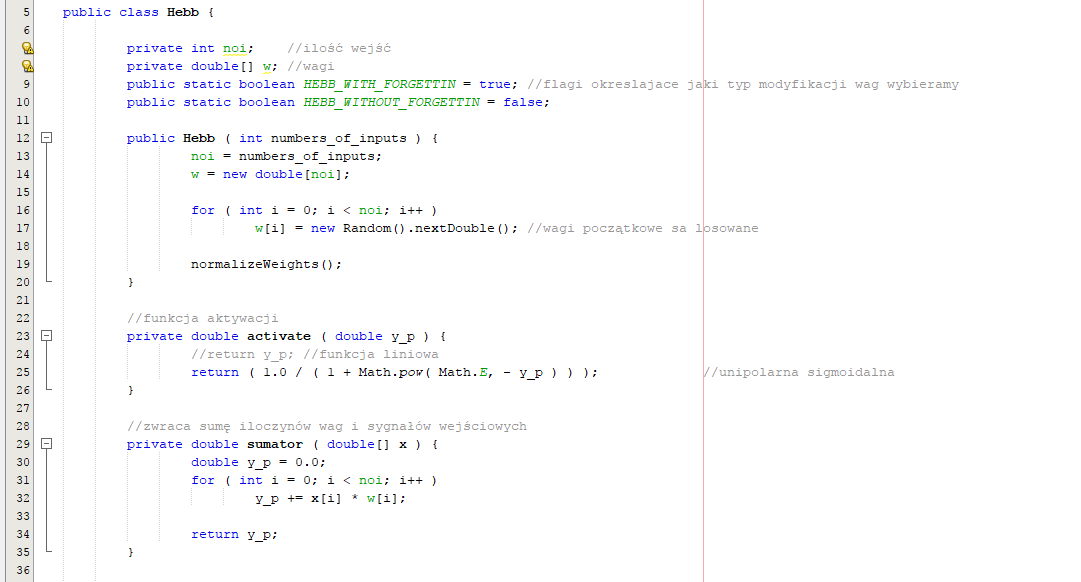
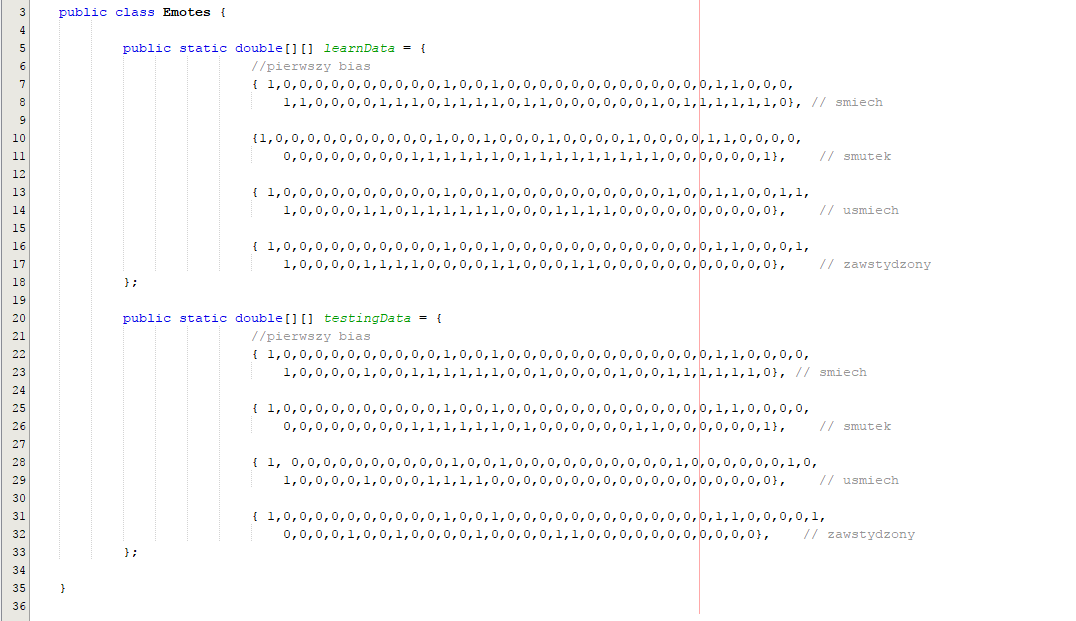
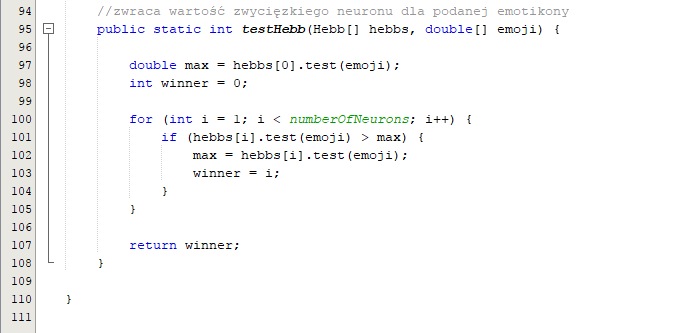
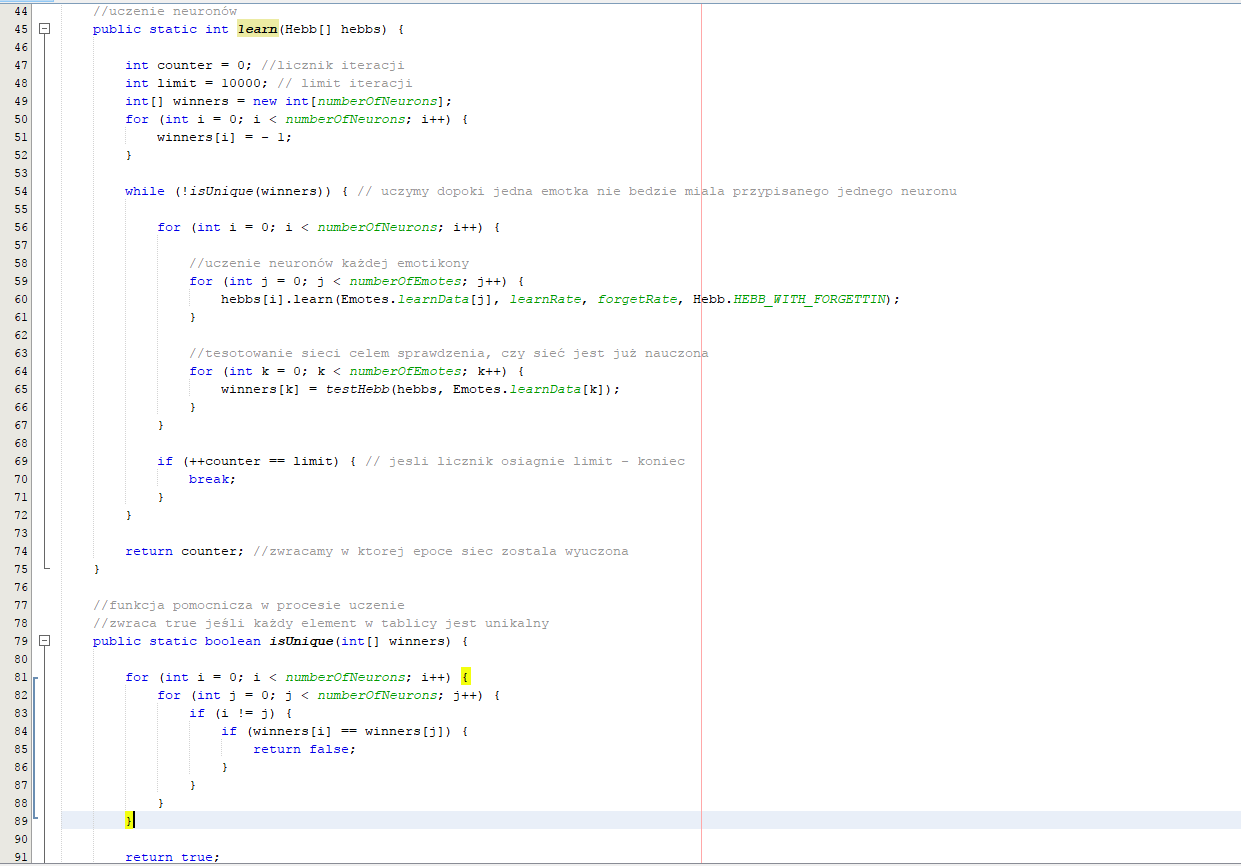
Na podstawie otrzymanych wyników pierwszy wniosek jaki się nasuwa to taki, że współczynnik zapominania pozwala osiągnąć wyniki na lepszym poziomie. Dla moich danych testujących sieć potrafiła osiągnąć 55% poprawnych wyników gdzie dla wariantu bez współczynnika zapominania było to 45%. Ważne jest jednak aby współczynnik zapominania nie był zbyt duży gdyż możemy doprowadzić do sytuacji gdzie sieć będzie zbyt szybko zapominać tego czego się nauczyła.

Wraz ze wzrostem współczynnika uczenia liczba epok potrzebnych do nauczenia sieci malała co wynika z faktu, że sieć szybciej dochodzi do rozwiązania dla danych uczących. Jednak testując ją podobnymi emotikonami myliła się ona co raz bardziej dla większych współczynników uczenia.

Błędy przy testowaniu sieci mogły wynikać z faktu, że emotikony testujące były różne od uczących w dosyć dużym stopniu, co mogło wpływać na to, że były one klasyfikowane do innych emotek niż powinny.

W uzyskaniu zadowalających wyników, ważne było znalezienie optymalnej wartości współczynnika uczenia i zapominania, które dla mojej sieci wynosiły odpowiednio 0,01 i 0,0033.

1. **Listing kodu.**

****

1. **Bibliografia**

Stanisław Osowski – Sieci neuronowe do przetwarzania informacji, ISBN 83-7207-615-4

<http://ecee.colorado.edu/~ecen4831/Demuth/Ch7_pres.pdf>

<http://pracownik.kul.pl/files/31717/public/Model_neuronu_Hebba.pdf>