|  |  |  |  |
| --- | --- | --- | --- |
| Sprawozdanie z ćwiczenia nr 2 | | | |
| *Nazwisko i imię* | *Nr grupy* | *Data* | *Ocena* |
| Robert Wróblewski | 2 | 27. lutego 2013 10:15 |  |
| Tomasz Cudziło | 2 |  |

# Przebieg wykonania ćwiczenia

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Wariant ustawień parametrów sieci neuronowej | | | | | | | | | | |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| Liczba epok uczących | 1000 | 500 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |
| Funkcja aktywacji Sigm/lin-prog [S/L] w warstwie wyjściowej | S | S | S | L | S | S | S | S | S | S | S |
| Liczba neuronów w 1. warstwie ukrytej | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | \* | 15 |
| Liczba neuronów w 2. warstwie ukrytej | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 15 | 0 | \* | 0 |
| Algorytm genetyczny  [T/N, populacja] | N | N | N | N | N | N | N | N | N | N | T |
| Zaszumienie sieci neuronowej po 100 epokach [T/N, wielkość] | N | N | N | N | T ,0.1 | N | T, 0.1 | N | N | N | N |
| Tasowanie faktów co 100 epok | N | N | N | N | N | T | T | N | T, 10 | N | N |

*Komentarz do wariantu nr 10:* Wykorzystano 5 warstw ukrytych składających się kolejno z 15 sigm., 14 sigm., 13 lin., 11 sigm. i 8 lin. neuronów.

# Analiza uzyskanych wyników

## Analiza statystyczna

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Wariant ustawień parametrów sieci neuronowej | | | | | | | | | | |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| Średni błąd względny [%] | 5,17% | 6,48% | 5,77% | 10,19% | 5,01% | 4,46% | 5,50% | 5,44% | 6,40% | 6,06% | 5,18% |
| Współczynnik korelacji | 0,91 | 0,95 | 0,89 | 0,86 | 0,94 | 0,95 | 0,96 | 0,92 | 0,94 | 0,90 | 0,94 |
| Błąd średniokwadratowy | 24783,93 | 31424,96 | 26951,97 | 42205,21 | 23323,29 | 22918,49 | 23941,88 | 25153,94 | 33867,07 | 26450,28 | 22330,81 |
| Błąd maksymalny [%] | 13,47% | 14,37% | 12,97% | 28,13% | 13,69% | 12,52% | 11,05% | 12,73% | 15,05% | 12,01% | 9,63% |

Zależność współczynnika korelacji od średniego błędu względnego kolejnych wariantów ustawień sieci neuronowych jest zobrazowana na wykresie ze strony 3.

Rysunek 1 Zależność współczynnika korelacji od średniego błędu względnego wyników uzyskanych przez kolejne warianty ustawień parametrów sieci neuronowej w stosunku do wartości ze zbioru danych testowych.

## Wnioski do zbiorczej analizy statystycznej

Na podstawie dokonanych testów nie jesteśmy w stanie jednoznacznie wyłonić najlepszego zestawu parametrów sieci neuronowej. Najniższy średni błąd względny równy 4,46% uzyskaliśmy dla wariantu nr 6. Jest to wariant, w którym w porównaniu do ustawień wyjściowych ustawiono tasowanie faktów co 100 epok. Należy też zwrócić uwagę na wariant nr 7, który posiadał jednocześnie ustawione tasowanie faktów i zaszumienie sieci. Uzyskano przy nim najwyższy współczynnik korelacji równy 0,96 jednak z gorszym średnim błędem względnym.

## Wnioski do analizy statystycznej kolejnych wariantów

### Wariant nr 1

Jest to zestaw wyjściowy parametrów sieci. Do niego porównujemy resztę przypadków.

### Wariant nr 2 – Zmniejszona liczba epok

Można zauważyć wzrost współczynnika korelacji w przypadku zmniejszenia ilości epok uczących z 1000, do 500. Widoczny jest efekt przetrenowania sieci dla 1000 epok uczących, a przez to wyuczenie się rozwiązań ze zbioru uczącego „na pamięć”. Niestety w związku z ograniczonym czasem zajęć laboratoryjnych nie było, bez wysoce negatywnego wpływu na ilość przetestowanych wariantów, możliwości przeprowadzenia wielu prób na pojedynczym wariancie współczynników sieci, w związku z czym nie możemy z całkowitą pewnością wykluczyć zaistnienia anomalii związanej z bardzo korzystnym dla dalszej nauki losowaniem początkowych wag neuronów w sieci.

### Wariant nr 3 – Dodatkowa warstwa ukryta

Dodanie drugiej warstwy ukrytej złożonej z 8 neuronów wpłynęło negatywnie na uzyskane rezultaty i czas nauki sieci. Brak poprawy był w tym wypadku w pełni zgodny z wcześniejszymi przewidywaniami. Dla względnie prostych problemów, jak omawiany, w pełni wystarczająca jest sieć z jedną warstwą ukrytą.

### Wariant nr 4 – Zmiana funkcji aktywacji w warstwie wyjściowej

Jako podstawową przyjęliśmy sigmoidalną funkcję aktywacji. Test dla sieci z liniową funkcją aktywacji w warstwie wyjściowej wykazał się gorszymi wynikami. Zastosowanie liniowej funkcji aktywacji negatywnie wpłynęło na uzyskane wyniki. Wariant nr 4 okazał się najgorszym z przetestowanych. Jak niefortunną decyzją byłoby wykorzystanie takiej właśnie sieci do rzeczywistego prognozowania wyraźnie ilustruje pierwszy wykres. Wariant 4, oznaczony ‘x’, znacznie odstaje od pozostałych wyników.

### Wariant nr 5 – Zaszumienie sieci neuronowej po 100 epokach