

Ćwiczenie opracował: dr hab. inż. Paweł Piotrowski, prof. uczelni  
Wersja (27.04.2022)

## ĆWICZENIE B4

### ZADANIE 1

#### SORTOWANIE KŁOCKÓW O RÓŻNYCH KSZTAŁTACH I KOLORACH SIECIĄ NEURONOWĄ KOHONENA

**Dane:** klocki o 2 różnych kolorach (zmienna jakościowa wejściowa) oraz o 2 wyraźnie odmiennych kształtach (opisane 3 liczbami – zmienne ilościowe wejściowe).

**Cel:** weryfikacja skuteczności grupowania klocków w zależności od sugerowanej liczby grup oraz weryfikacja kształtu ułożenia neuronów wyjściowych oraz włączenia opcji sąsiedztwa oraz wyłączenia opcji sąsiedztwa w procesie uczenia sieci neuronowej Kohonena

**Możliwe grupy:** 2 kategorie (2 neurony wyjściowe), podział na 4 kategorie (4 neurony wyjściowe)

**Narzędzia:** Sieć neuronowa samoorganizująca Kohonena – pakiet Statistica, arkusz kalkulacyjny.

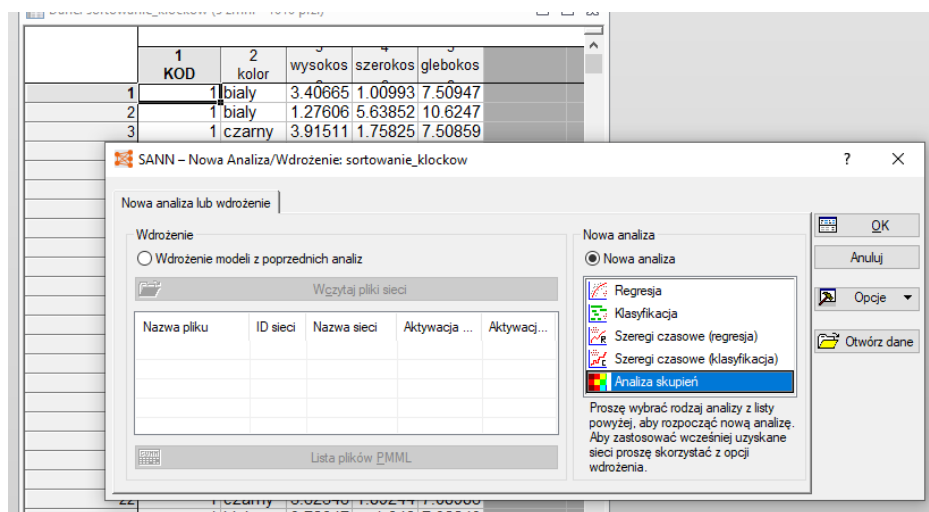
#### Instrukcja wykonania ćwiczenia.

w programie Statistica: Menu -> nowy -> arkusz (podajemy: 1002 przypadki, 5 zmiennych)

w arkuszu kalkulacyjnym zaznaczamy zakres 11;E1004 czyli z nagłówkami całość danych z zakładki "kolorowe\_klocki"

w statistica wklejamy do arkusza -> wklej z nagłówkami -> wklej z nazwami zmiennych

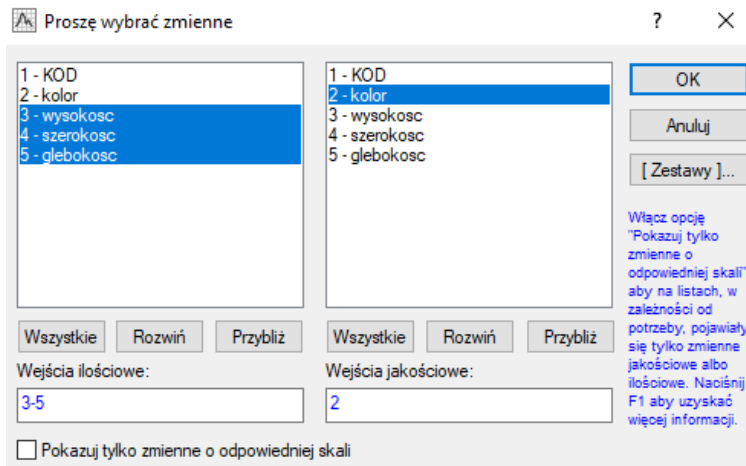
menu -> Data Mining -> Sieci neuronowe -> nowa analiza -> wybieramy: analiza skupień



menu podstawowe –klikamy ->zmiennie

wejścia ilościowe – wpisujemy 3-5 (wysokość, szerokość, głębokość)

wejścia jakościowe – podajemy 2 (kolor)



menu -> wybór podzbiorów klikamy „identyfikator próby”

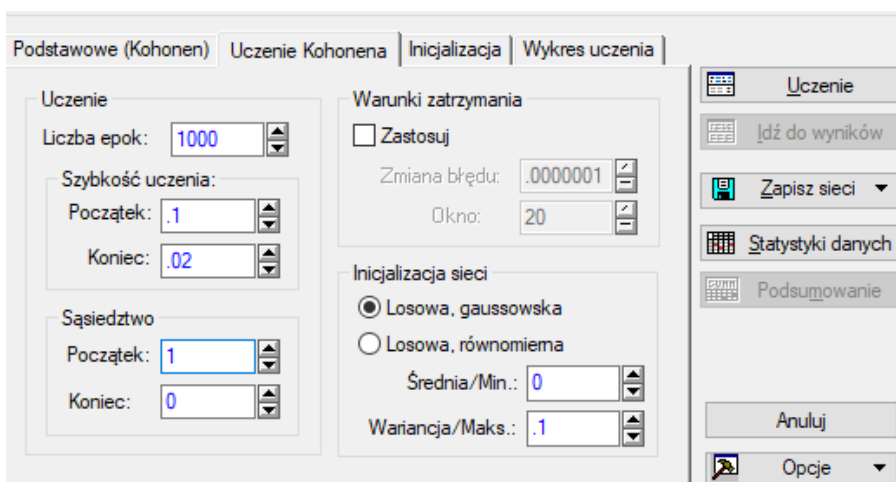
klikamy w pole "ucząca" - klikamy w "stan" włączona - klikamy w pole "zmienna identyfikująca próby" wybieramy zmienną "KOD", w polu "kod próby uczącej" wpisujemy "1", klikamy ok.

klikamy w pole "testowa" - klikamy w "stan" włączona - klikamy w pole "zmienna identyfikująca próby" wybieramy zmienną "KOD", w polu "kod próby testowej" wpisujemy "2", klikamy ok.

klikamy OK. w głównym oknie w prawym górnym rogu

w menu „podstawowe kohonen” wybieramy liczbę neuronów wyjściowych zgrupowanych w postaci szeregu (nie obszaru) tzn. wysokość 1 a szerokość 2 (grupowanie na 2 kategorie) lub 2x2 lub 1x4 (grupowanie na 4 kategorie)

Menu „uczenie kohonena” – możliwość modyfikacji parametrów np. liczba epok uczących, rodzaj inicjalizacji sieci wybór czy ma być sąsiedztwo i jakie sąsiedztwo



Menu „wykresy uczenia” – można zaznaczyć wizualizację uczenia.

Klikamy UCZENIE w zakładce podstawowe kohonen

Po zakończeniu nauki pojawi się okno SANN wyniki

Klikamy w zakładkę „wykres kohonena” – widać wizualizację grup

Klikamy w zakładkę „predykcja (kohonen)”

Klikamy w licznosci (lewe) widać wizualizację przypisania do grup

Klikamy w licznosci (prawe) widać wizualizację liczebności grup

W oknie głównym SANN wyniki zaznaczamy pole test tuż pod polem uczenie aby zobaczyć wyniki dla zakresu uczenia i testowania

W oknie głównym odznaczamy „aktywacja neuronu zwycięskiego” a zaznaczamy „neuron zwycięski”

Klikamy w pole „predykcja”

Kolumna „pozycja” podaje położenie neuronu zwycięscy dla danego profilu a kolumna ID neuronu unikalny numer.

Kopiujemy kolumnę ID neuronu i wklejamy do pliku excela – zakładka „kolorowe\_klocki” kolumna G (pierwszy wariant grupowania na 2 kategorie)

Powtarzamy badanie dla grupowania na 4 kategorie czyli 4 neurony wyjściowe (należy od nowa sieć Kohonena zbudować).

Wykonujemy obliczenia w arkuszu kalkulacyjnym odpowiednie aby zweryfikować skuteczność grupowania w podziale na 2 grupy oraz 4 grupy w zakresie danych treningowych oraz testowym. Zweryfikować czy sąsiedztwo oraz kształt położenia neuronów wyjściowych ma wpływ na jakość grupowania.

## ZADANIE 2

### GRUPOWANIE PROFILI OBCIĄŻENIA KSE SIECIĄ NEURONOWĄ KOHONENA

**Dane:** miesiące styczeń oraz luty z lat 2002-2004 (udostępnione w pliku – dane\_kse.xls)

**Cel:** podział profili na grupy - określenie właściwej liczby grup na podstawie procentu przypisań profili do każdej grupy oraz ustalenie skuteczności przypisania do grupy przez sieć neuronową Kohonena.

**Możliwe grupy:**

3 grupy (dzień roboczy, sobota, niedziela)

2 grupy (dzień roboczy, dzień świąteczny).

**Narzędzia:** Sieć neuronowa samoorganizująca Kohonena – pakiet Statistica, arkusz kalkulacyjny.

### **Instrukcja wykonania ćwiczenia.**

W arkuszu kalkulacyjnym w zakładce „wykresy\_typy\_dni” wykonujemy 3 wykresy liniowe – osobno dla profili dni roboczych (1), sobót (2) oraz niedziel (3) stosując autofiltr dla danych w zakładce „dane\_kse” – obserwujemy różnice pomiędzy zgrupowanymi 3 typami profili. Wynik obserwacji umieszczamy w sprawozdaniu w punkcie „wnioski”.

w programie Statistica: Menu -> nowy -> arkusz (podajemy: 175 przypadków, 25 zmiennych)

w arkuszu kalkulacyjnym zaznaczamy zakres E1;AC176 czyli z nagłówkami całość danych z zakładki "Dane\_kse"

w statistica wklejamy do arkusza -> wklej z nagłówkami -> wklej z nazwami zmiennych

menu -> Data Mining -> Sieci neuronowe -> nowa analiza -> wybieramy: analiza skupień

menu podstawowe –klikamy ->zmiennie

wejścia ilościowe – wpisujemy 2-25 lub zaznaczamy od H1 do H24

menu ->wybór podzbiorów klikamy „identyfikator próby”

klikamy w pole "ucząca" - klikamy w "stan" włączona - klikamy w pole "zmienna identyfikująca próby" wybieramy zmienną "kod", w polu "kod próby uczącej" wpisujemy "1", klikamy ok.

klikamy w pole "testowa" - klikamy w "stan" włączona - klikamy w pole "zmienna identyfikująca próby" wybieramy zmienną "kod", w polu "kod próby testowej" wpisujemy "2", klikamy ok.

klikamy OK. w głównym oknie w prawym górnym rogu

w menu „podstawowe kohonen” wybieramy liczbę neuronów wyjściowych zgrupowanych w postaci szeregu (nie obszaru) tzn. wysokość 1 a szerokość 2 lub 3 w zależności od oczekiwanej liczby grup. W pierwszym kroku podajemy 3.

Menu „uczenie kohonena” – możliwość modyfikacji parametrów np. liczba epok uczących, rodzaj inicjalizacji sieci

Menu „wykresy uczenia” – można zaznaczyć wizualizację uczenia.

Klikamy UCZENIE w zakładce podstawowe kohonen

Po zakończeniu nauki pojawi się okno SANN wyniki

Klikamy w zakładkę „wykres kohonena” – widać wizualizację grup

Klikamy w zakładkę „predykcja (kohonen)

Klikamy w liczności (lewe) widać wizualizację przypisania do grup

Klikamy w liczności (prawe) widać wizualizację liczebności grup

W oknie głównym SANN wyniki zaznaczamy pole test tuż pod polem uczenie aby zobaczyć wyniki dla zakresu uczenia i testowania

W oknie głównym odznaczamy „aktywacja neuronu zwycięskiego” a zaznaczamy „neuron zwycięski”

Klikamy w pole „predykcja”

Kolumna „pozycja” podaje położenie neuronu zwycięscy dla danego profilu a kolumna ID neuronu unikalny numer.

Kopiuujemy kolumnę ID neuronu i wklejamy do pliku excela – zakładka „wyniki” kolumna F

Wykorzystując autofiltr liczymy liczebności w zakładce „wyniki\_Statistica”,  
Liczebności liczymy osobno dla zakresu uczenia (1) oraz zakresy testowania (2)  
Gdy w autofiltrze ustawimy zakres (1) to wtedy wybieramy w kolumnie C wartość 1 a w kolumnie F wybieramy 1, potem 2 a potem 3 i notujemy ile razy wystąpiły w kolumnie F kolejno jedynki, dwójki i trójki. Potem wybieramy w kolumnie C liczbę 2 i powtarzamy liczenie, następnie w kolumnie C wybieramy liczbę 3 i powtarzamy liczenie.

Następnie uzyskane dane liczbowe wstawiamy do zakładki „wyniki\_grupowanie” – w pierwszym badaniu do komórek C7:E9

Powtarzamy badanie dla 2 grup (należy od nowa sieć Kohonena zbudować).

Ewentualnie zmieniamy parametry sieci Kohonena i badamy wpływ parametrów na wyniki grupowania dla jednego grupowania (podział na 2 lub 3 grupy).

Dla zakresu uczenia wykonać w arkuszu kalkulacyjnym trzy wykresy (podział na 3 grupy). Dane do wykresu to profile przypisane do danego ID neuronu (z wszystkich trzech typów dni) przez sieć neuronową Kohonena – we wnioskach opisać wynik obserwacji różnic pomiędzy 3 grupami. Najwygodniej wyniki kolumny F z zakładki „wyniki\_Statistica” skopiować do zakładki „dane\_kse” do kolumny AD i następnie włączyć autofiltr, ustawić odpowiednią wizualizację danych i wybrane dane skopiować do zakładki „wykresy\_grup\_ID”. Następnie wykonać wykres na podstawie danych. Czynności powtórzyć dwukrotnie aby otrzymać łącznie trzy wykresy-każdy dla osobnego ID neuronu.

### **Sprawozdanie.**

1. Wstęp teoretyczny – sieć Kohonena
2. Wyniki tabelarycznie ze Statistica
3. Wyniki przetworzone (tabele) – podział na grupy, liczebności grup, liczebności grup procentowo, opis parametrów sieci Kohonena.
4. Wybór najlepszego grupowania – najwięcej prawidłowych przypisań do grupy.
5. Wnioski.