

# Kolejność prezentacji danych uczących a skuteczność procesu nauki w sieciach MLP

Cezary Dendek

Prof. nzw dr hab. Jacek Mańdziuk

# Plan prezentacji

- Czy kolejność prezentacji wzorców wpływa na przebieg procesu uczenia?
- Prezentacja metody porządkowania
- Badane zagadnienie
- Wyniki
- Przypuszczenia
- Dyskusja

# Czy kolejność prezentacji wzorców wpływa na przebieg procesu uczenia?

- Przykłady
  - Sieci Kohonena  
długie ciągi uczące, niewielka ilość neuronów
  - MLP  
uzyskane wyniki:
    - 2-3 krotne różnice RMSE
    - znaczna różnica w jakości działania sieci

# Metoda porządkowania

- Założenia
  - Grupowanie elementów odległych od siebie na jednym końcu ciągu, bliskich sobie – na drugim
  - Wykorzystanie metryki określonej na przestrzeni, z której pochodzą elementy zbioru

# Metoda porządkowania

- Metryka

- Określona na całym wzorcu (wejście i wyjście) – na poszczególnych podprzestrzeniach

- Normalizowana:

$$M(x, y) = \sqrt{\sum_{i=1}^n \left( g_i \frac{d_i(x, y)}{\bar{d}_i} \right)^2}$$

- Ważona:

$$M(x, y) = \sqrt{\sum_{i=1}^n g_i d_i(x, y)^2}$$

# Metoda porządkowania

- Metryka normalizowana
  - Postać zależy od danych wejściowych
  - Podprzestrzenie o znacznie się różniących wielkościach odległości są „centrowane” w punkcie 1
- Metryka ważona
  - Wagi arbitralnie przypisywane do poszczególnych podprzestrzeni

# Metoda porządkowania

- Ciąg „zgodny”  $\{T_k\}_{k=1}^K$

$$S_n = \frac{2}{(n-1)n} \sum_{k=1}^n \sum_{l=k+1}^n M(T_k, T_l)$$

Jest to wartość średnia odległości pomiędzy pierwszymi  $n$  elementami ciągu

# Metoda porządkowania

- Ciąg „zgodny”
  - Taki, aby  $\forall_{2 \leq k \leq K-1} S_k \geq S_{k+1}$
  - Jego postać zależy od metryki
  - Nie jest jednoznaczny
  - Nie zawsze istnieje
- Miary zgodności
  - Ilość inwersji
  - Ilość uogólnionych inwersji



# Metoda porządkowania

- Znajdowanie ciągu „zgodnego”
  - Tablica odległości  $D$  elementu od pozostałych elementów
  - Wybór elementu o najmniejszej wartości w  $D$  i dołączenie go na początku ciągu
  - Aktualizacja  $D$
  - $O(n^2)$

# Prezentacja zagadnienia

- Rozpoznawanie cyfr pisanych odręcznie
- Wejście – obraz 17x19
- Wyjście – wektor 1x10
- Wielkość zbioru – 6000 elementów
- Zbiór uczący  $T$ ,  $\#T=5500$
- Zbiór walidacyjny  $V$ ,  $\#V=500$

# Metody

- Sieć:
  - 323 wejścia
  - {20, 30, 35} sigmoidalnych neuronów ukrytych
  - 10 wyjść

# Metody

- Metryka - podprzestrzenie
  - wejściowa z metryki Hamminga

$$G_1(v, w) = \min_{x, y \in \{-2, -1, 0, 1, 2\}} H(v, T(x, y, w)) + |x| + |y|$$

- wyjściowa: metryka dyskretna

# Metody

- Metryki użyte do porządkowania

- $M_1 = G_1$

- $\forall_{k,l \in O_2} \forall_{v,w \in O_1} M_2([v,k], [w,l]) = \sqrt{\frac{G_1(v,w)^2}{323^2} + G_2(k,l)}$

- 

- $\forall_{k,l \in O_2} \forall_{v,w \in O_1} M_3([v,k], [w,l]) = \sqrt{\left(\frac{G_1(v,w)}{G_1}\right)^2 + \left(\frac{G_2(k,l)}{G_2}\right)^2}$

# Metody

- Statystyki
  - RMSE
  - Ilość dobrze rozpoznanych elementów
  - Suma pobudzeń, dla których oczekiwano braku pobudzenia

$$IIS = \sum_{i=1}^N \sum_{j=1}^k o[i, j] I(d[i, j], 1)$$

$$I(a, b) = \begin{cases} 0 & \text{jeżeli } a=b \\ 1 & \text{jeżeli } a \neq b \end{cases}$$

# Metody

- Statystyki
  - Środek ciężkości błędu względem (permutowanego) ciągu uczącego

$$ECS = \frac{\sum_{i=1}^N i \sum_{j=1}^k e[i, j]}{\sum_{i=1}^N \sum_{j=1}^k e[i, j]}$$

# Wyniki

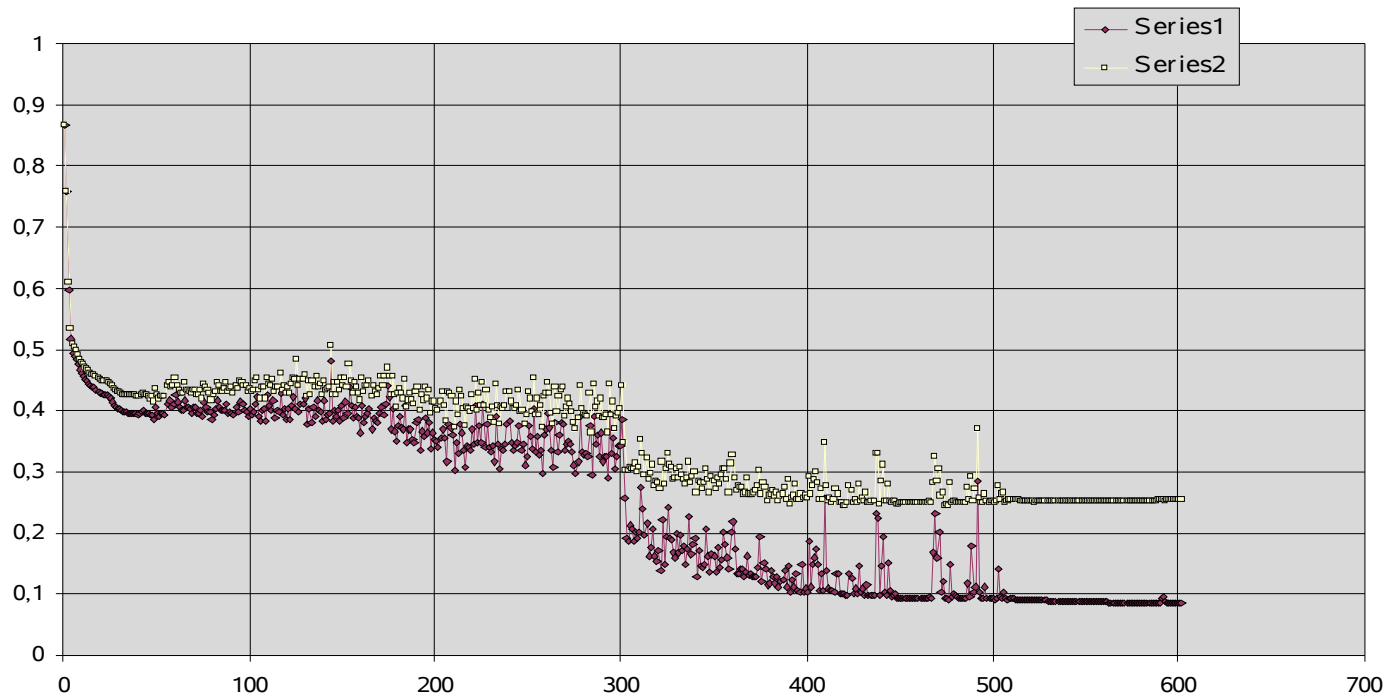
- Analizie (i magazynowaniu) podlegały odpowiedzi na poszczególne wzorce uczące na zbiorach
  - Uczącym w trakcie nauki
  - Uczącym – walidacja
  - Walidacyjnym (– walidacja)



# RMSE

- 30 neuronów, uczenie mieszane  
300xrO1+300xRND [inicjalizacja wag]

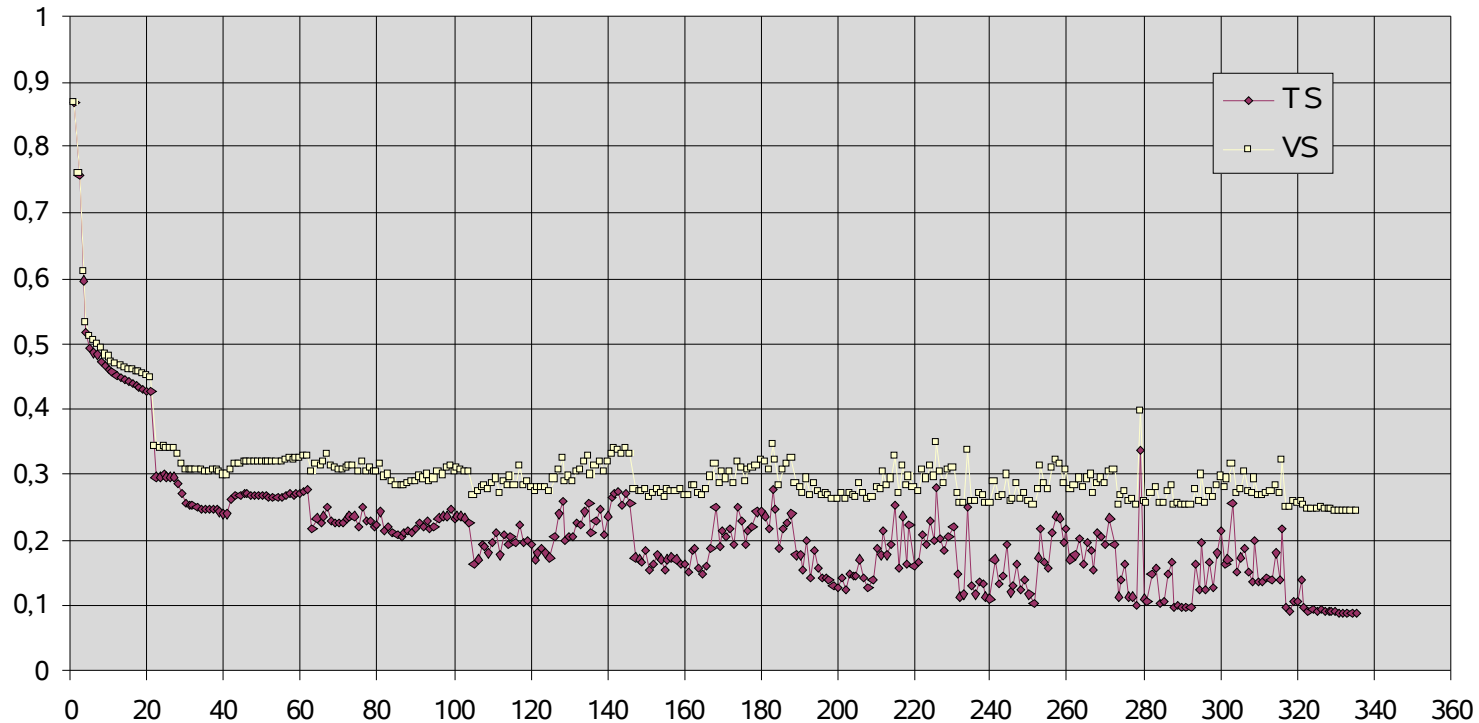
RMSE rO1x300+RND3x300



# RMSE

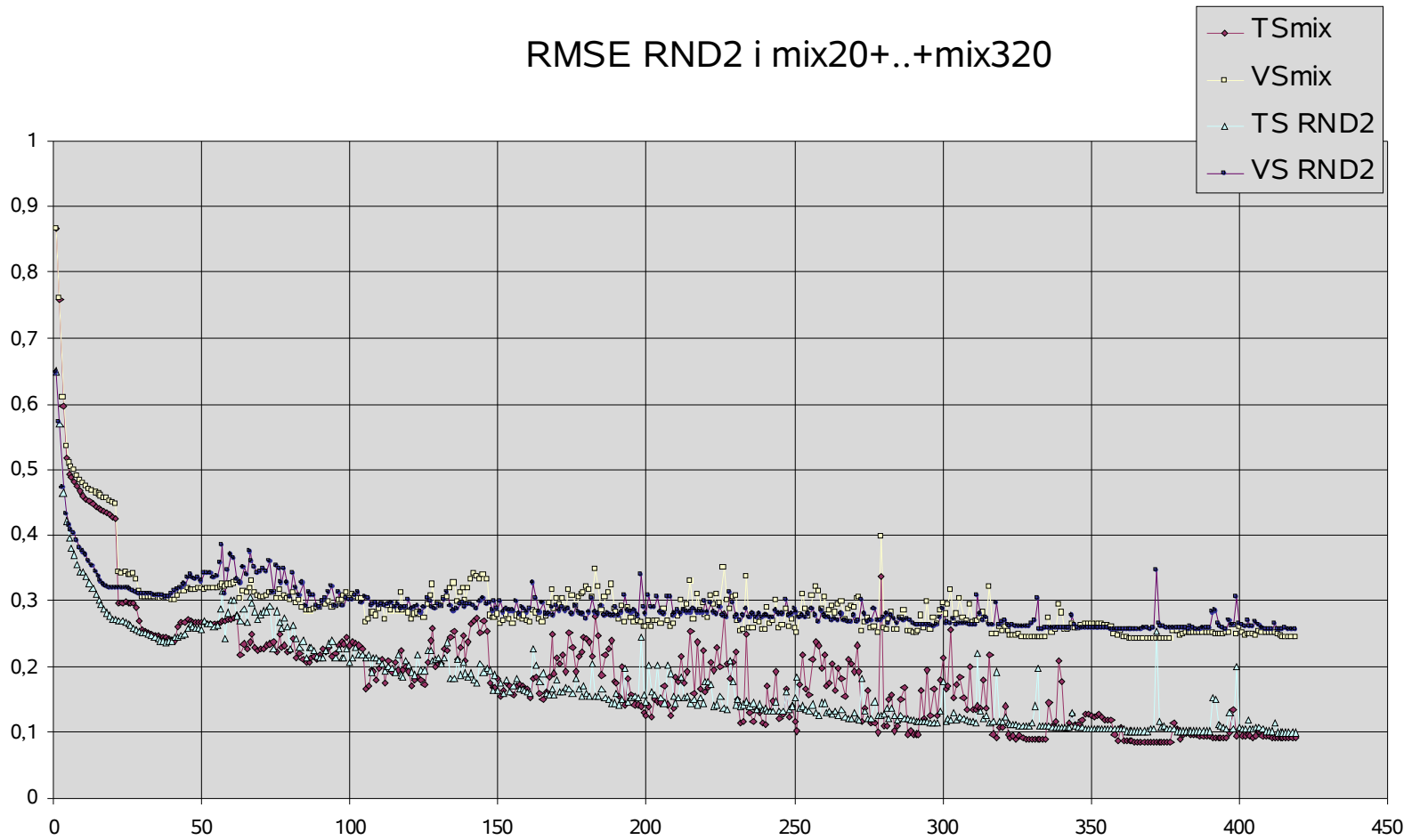
- 30 neuronów, uczenie mieszane (20xR1+20RND)xn

RMSE (rO1x20+RND2x20)xn



# RMSE

RMSE RND2 i mix20+..+mix320

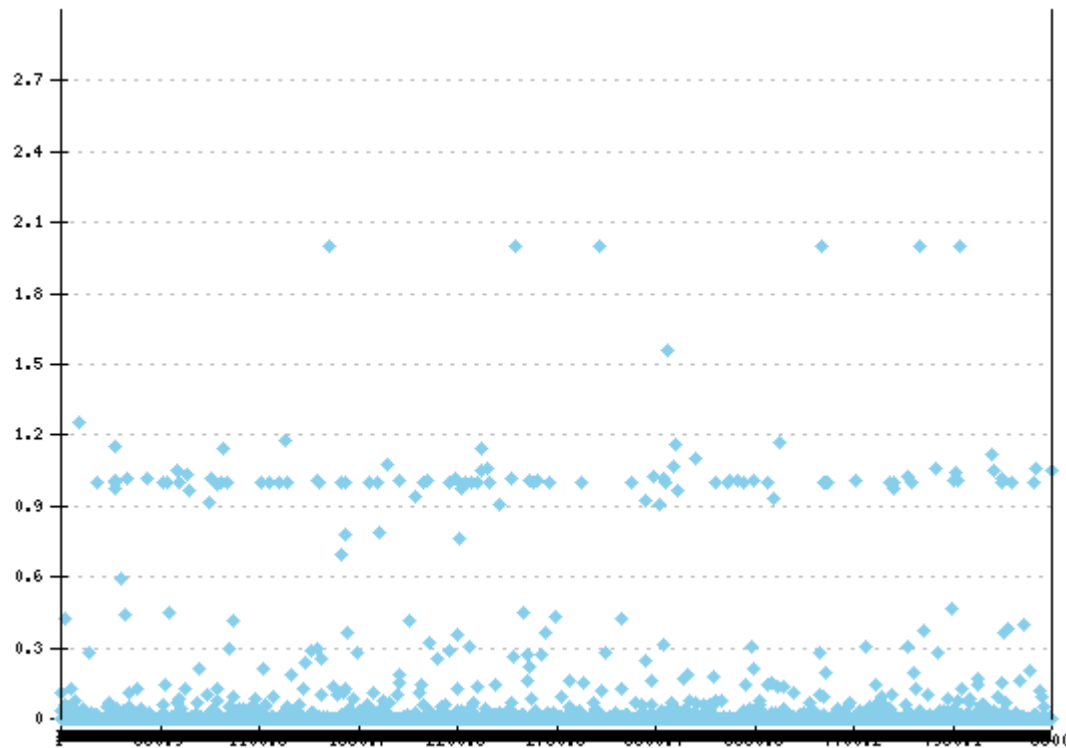


# RMSE

- Dla mieszanych ciągów losowych -  
wypłaszczenie

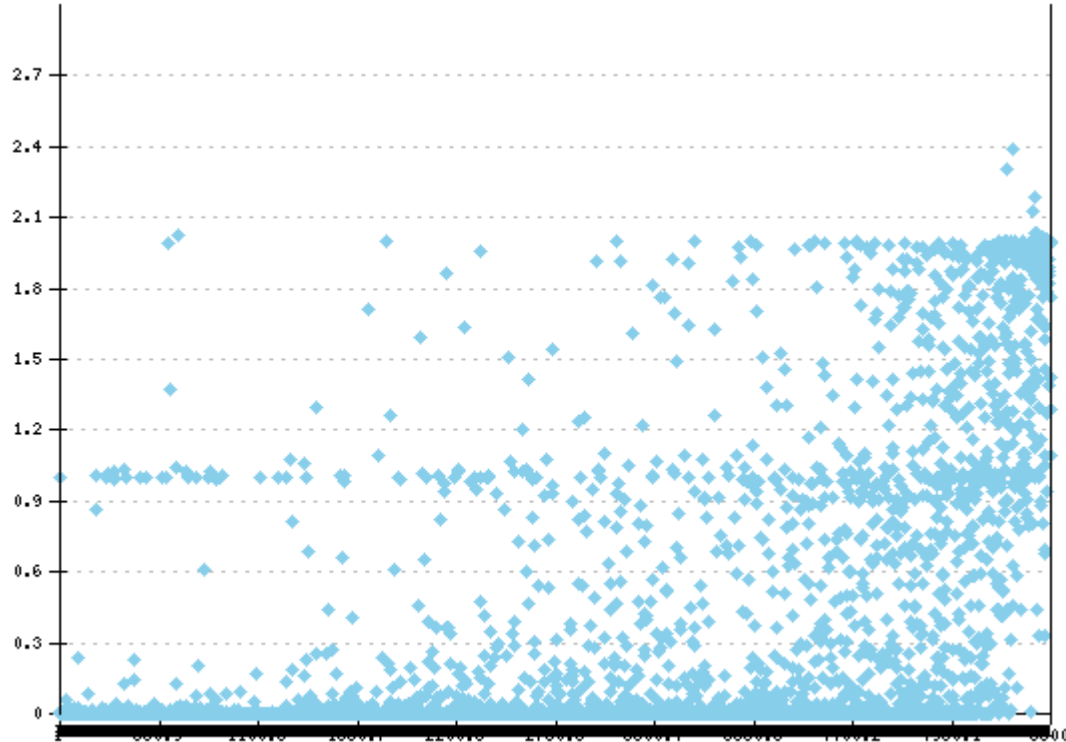
# Rozkład błędu na ciągu

- Ciąg losowy, uczący, 200 krok uczenia



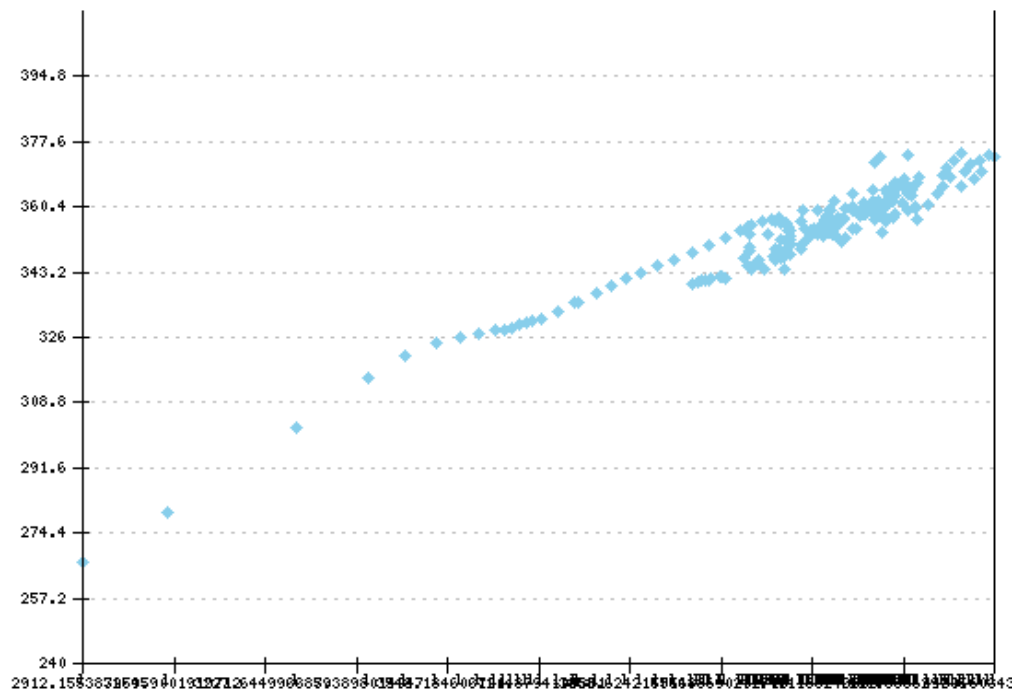
# Rozkład błędu na ciągu

- Ciąg odwrócony, uczący, 200 krok

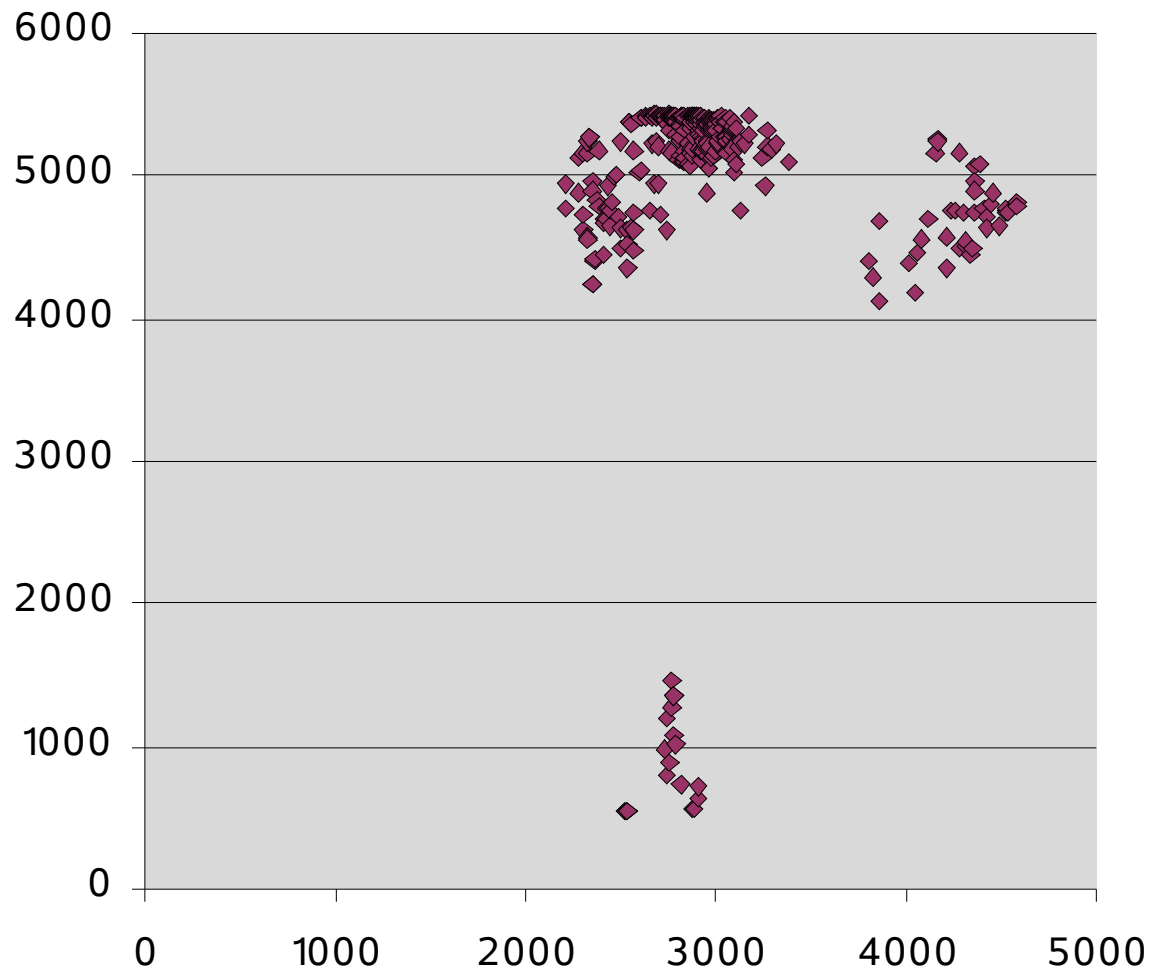


# Generalizacja ECS

- Walidacja na zbiorze uczącym i walidacyjnym ciąg rO1R, siec 30

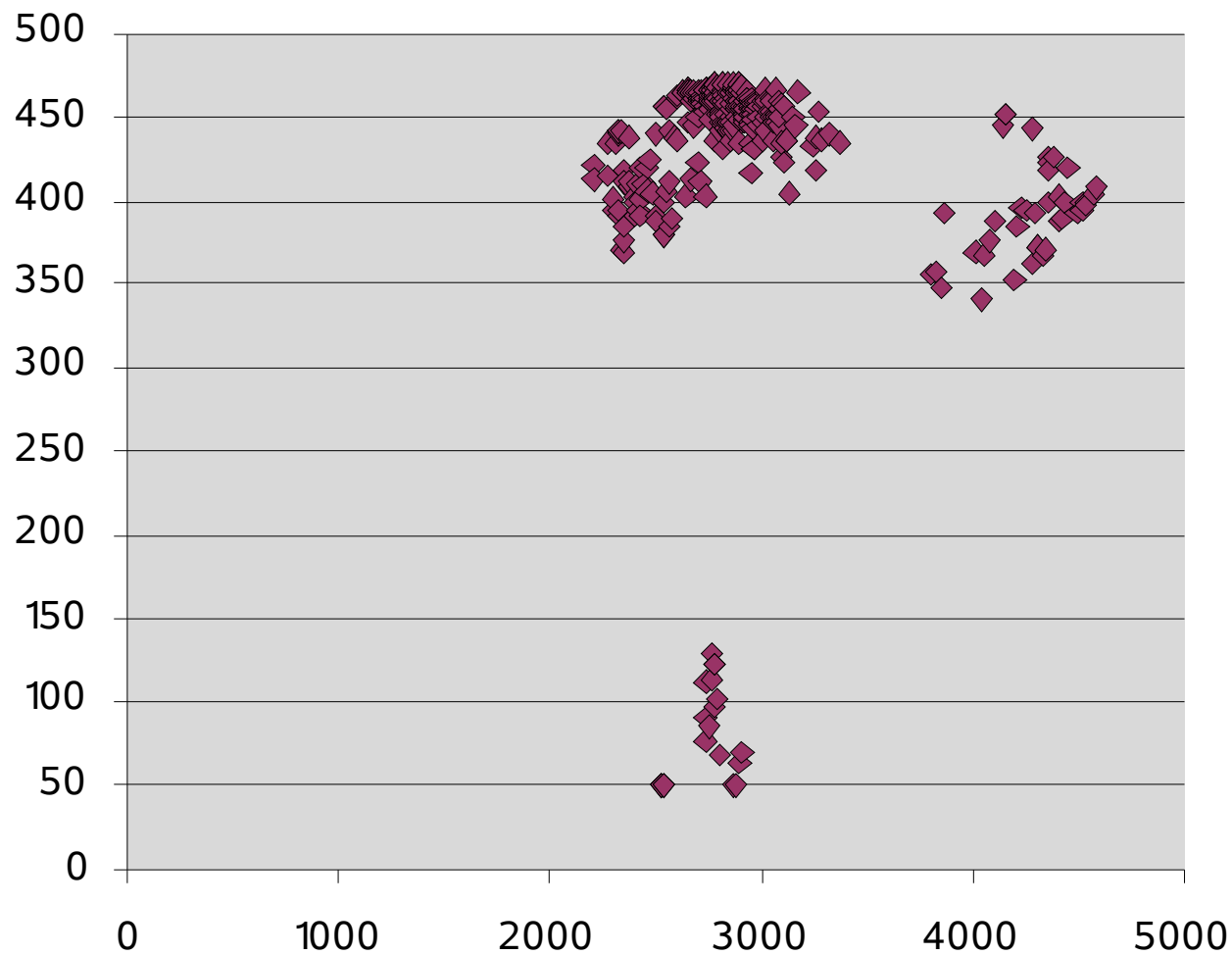


# ECS a PR – zbiór uczący

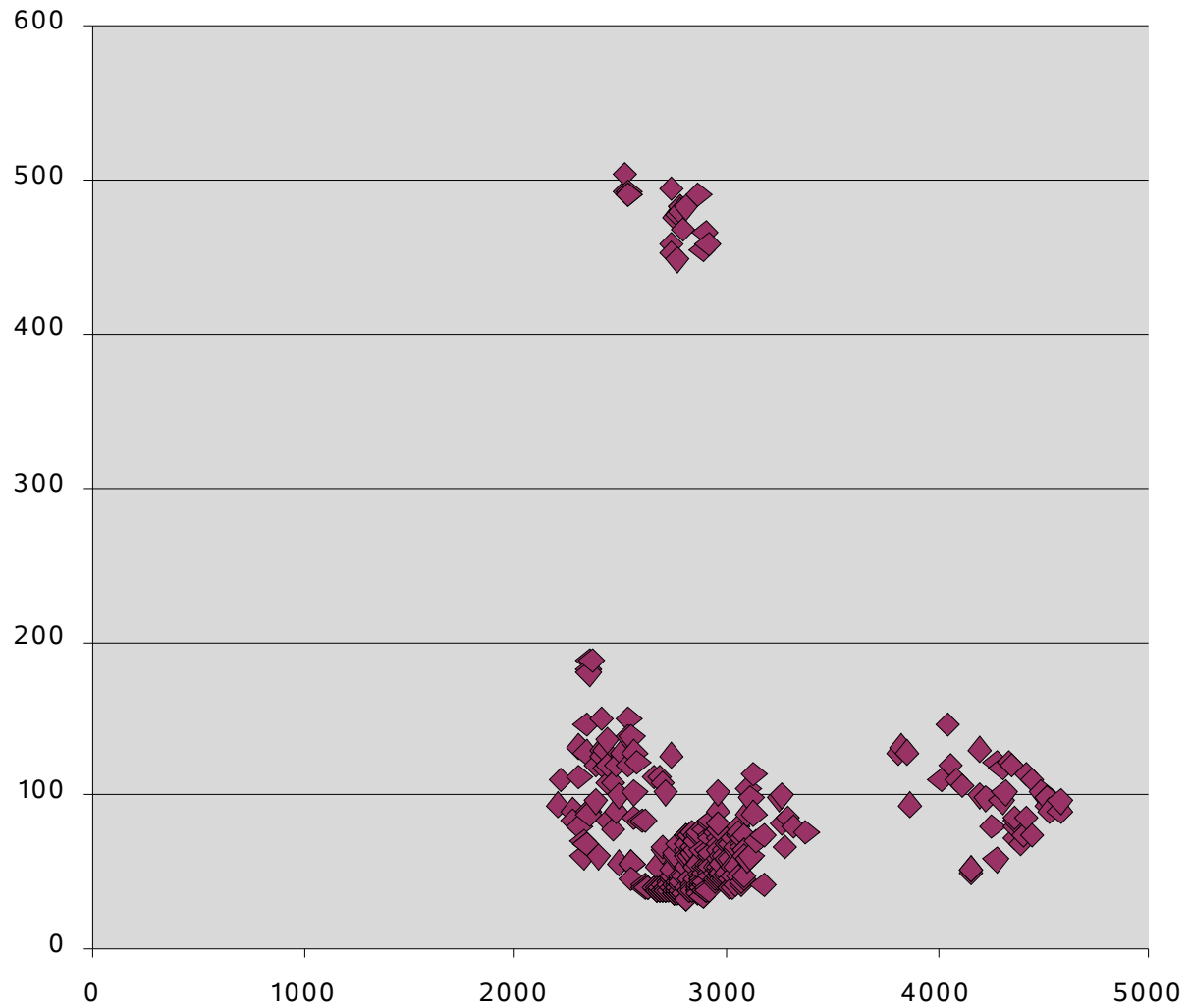




# ECS a PR - walidacja



# ECS a IIS



# Hipotezy

- Wyżarzanie
- Badanie rozstępu między centrami skupienia

Dziękuję  
za Państwa Uwagę