**Analiza**

W projekcie podjęłam temat klasyfikacji binarnej, wieloklasowej oraz regresję.

Rozwiązanie zadania klasyfikacji umożliwia szereg technik i algorytmów uczących. Takie algorytmy nazywamy klasyfikatorami, w analizie zostały użyte klasyfikatory oparte o : drzewa decyzyjne, sieci neuronowe, KNN. W przypadku każdego klasyfikatora została zrobiona kroswalidacja ( 8 razy), gdzie za każdym razem została wylosowana część danych uczących i testowych (walidacyjnych). Na końcu wyniki dla modeli o tych samych parametrach zostały uśrednione i zaprezentowane są w dalszej części raportu.

**KLASYFIKACJA WIELOKLASOWA**

Natomiast do klasyfikacji wieloklasowej został wykorzystany zbiór car. Bezpośrednio wiąże CAR z sześcioma atrybutami wejściowymi: zakup, konserwacja, drzwi, osoby, bezpieczeństwo i bagażnik.

Zbiór Car

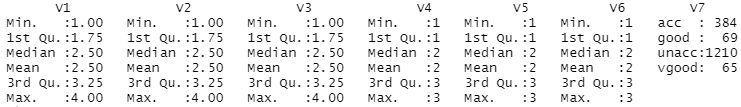
Wartości klas:  
  
unacc, acc, good, vgood  
  
Atrybuty

V1 = buying: vhigh, high, med, low.  
V2 = maint: vhigh, high, med, low.  
V3 = doors: 2, 3, 4, 5more.  
V4 = persons: 2, 4, more.  
V5 = lug\_boot: small, med, big.  
V6 = safety: low, med, high.

**Szczegóły**

W pierwszym etapie zaimportowano bazę oraz sprawdzono ewentualne braki. Zgodnie z opisem na <https://archive.ics.uci.edu/ml/machine-learning-databases/car/> dane są kompletne. Dane zostały zamienione na wartości numeryczne, ze względu na uproszczenie obliczeń różnych klasyfikatorów.

**Tabela – zbiór Car**



**KLASYFIKACJA BINARNA**

Dane pochodzą z

<https://archive.ics.uci.edu/ml/machine-learning-databases/00267/data_banknote_authentication>

1372 obserwacji oraz 5 zmiennych

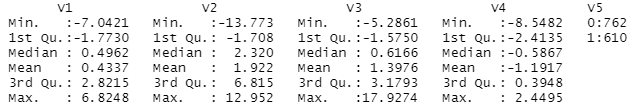
V1 = Variance

V2 = Skewness

V3 = Curtosis

V4 = Entropy

Wszystkie zmienne są numeryczne.



**REGRESJA**

Dane pochodzą z <https://archive.ics.uci.edu/ml/machine-learning-databases/servo>

167 obserwacji oraz 5 zmiennych

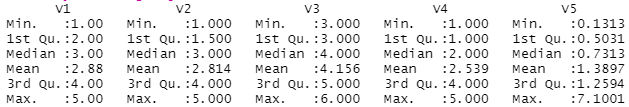
V1 = Motor

V2 = Screw

V3 = Pgain

V4 = Vgain

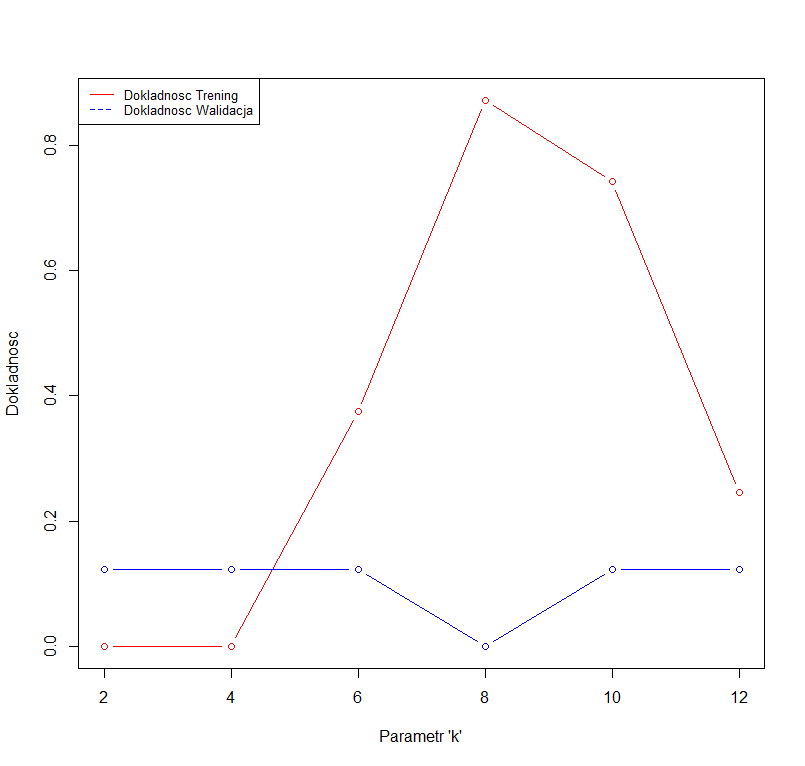
Sprawdzono braki danych, zbiór nie posiada takich obserwacji.



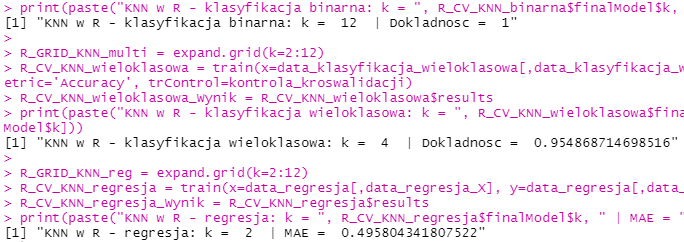
**KNN kod własny – Klasyfikacja binarna**

**Obraz zawierający tekst

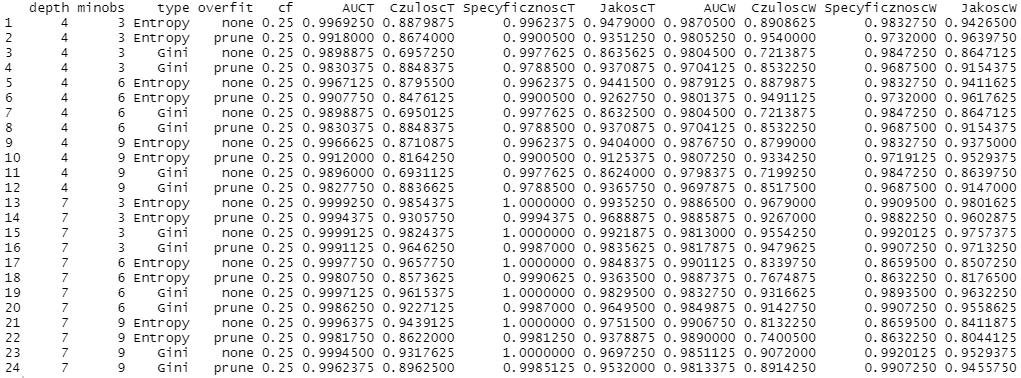
Opis wygenerowany automatycznie**

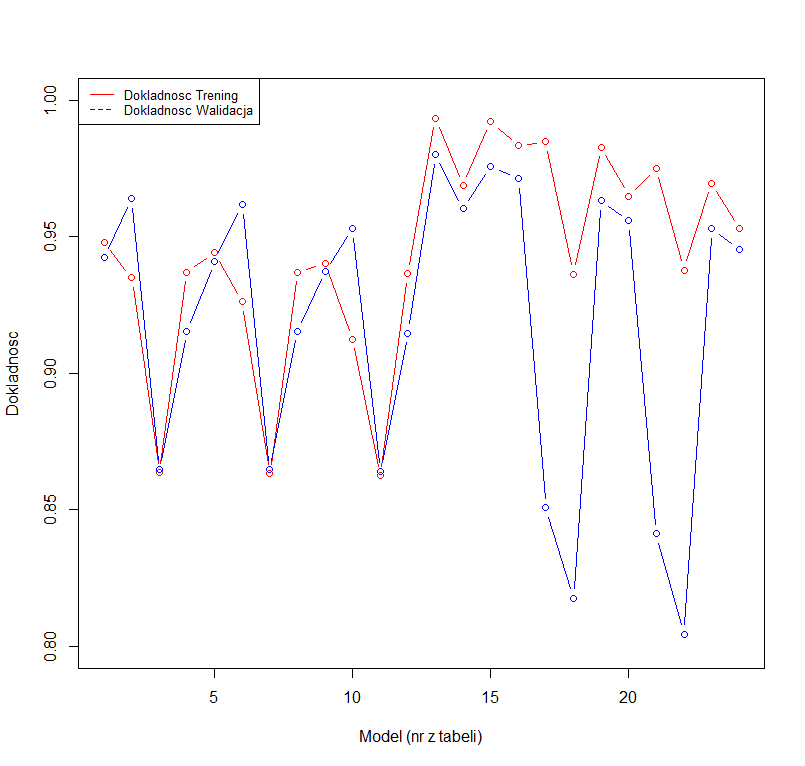
****

**PAKIET KNN – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

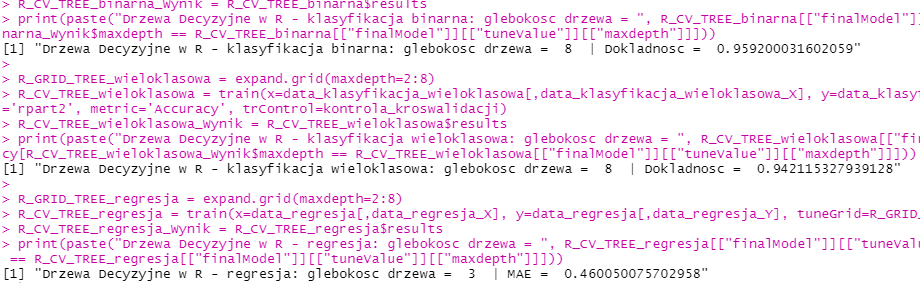
****

**DRZEWA DECYZYJNE kod własny – Klasyfikacja binarna**

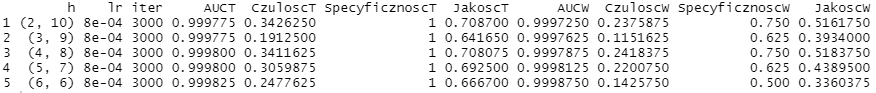
****

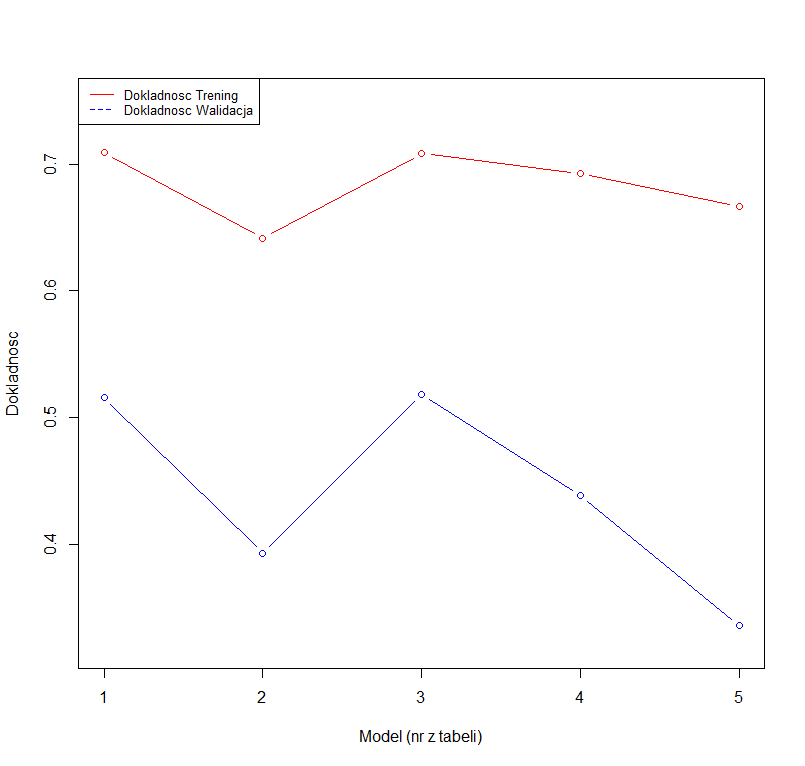
****

**PAKIET RPART – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

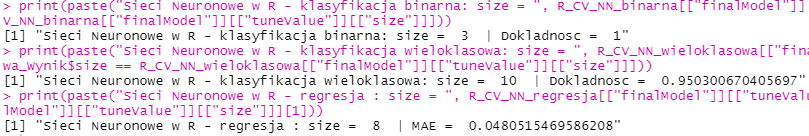
****

**SIECI NEURONOWE kod własny – Klasyfikacja binarna**

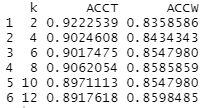
****

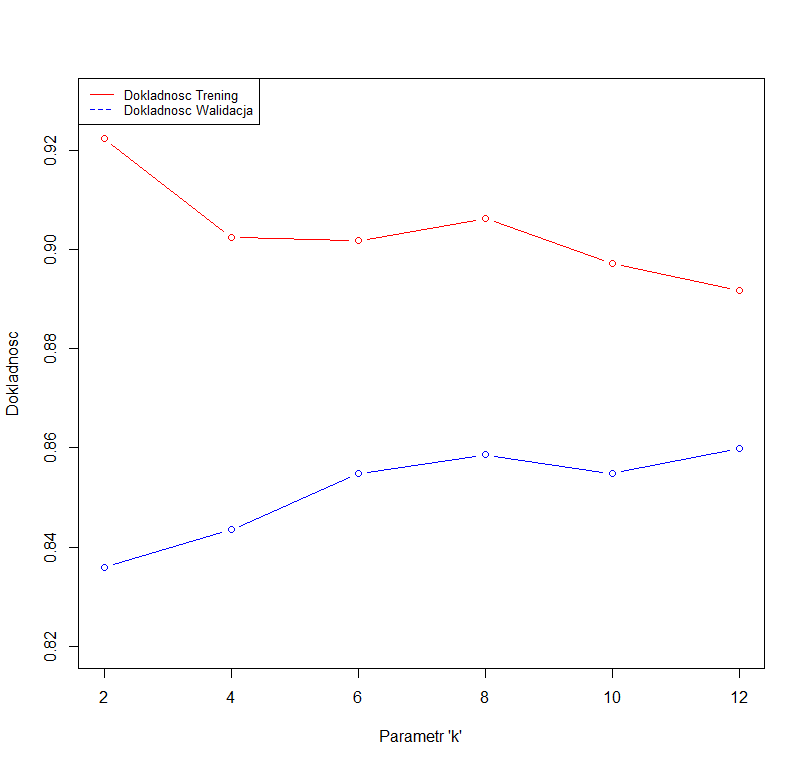
****

**PAKIET NNET – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

****

**KNN kod własny – Klasyfikacja wieloklasowa**

****

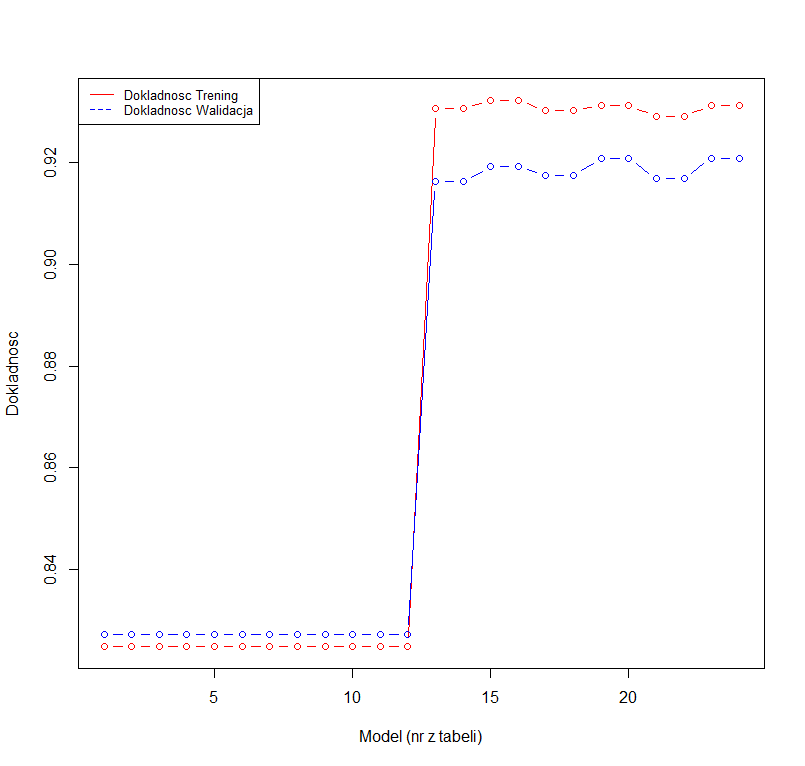
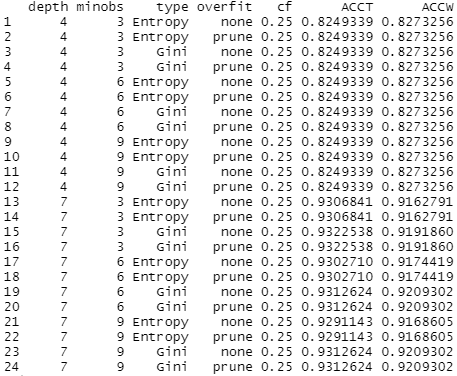
****

**PAKIET KNN – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**DRZEWA DECYZYJNE kod własny – Klasyfikacja wieloklasowa**

****

**PAKIET RPART – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

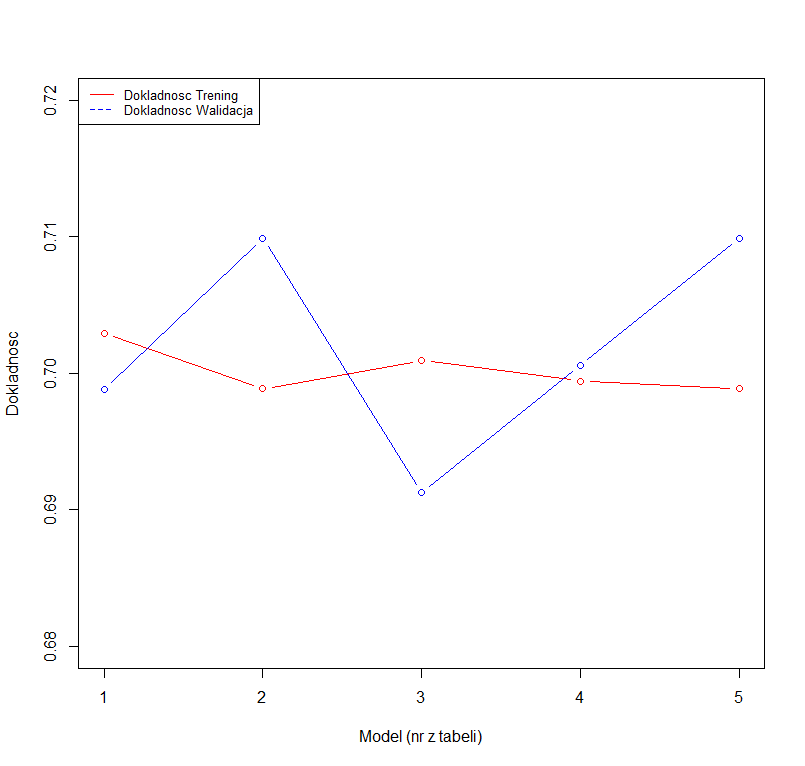
**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**SIECI NEURONOWE kod własny – Klasyfikacja wieloklasowa**

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

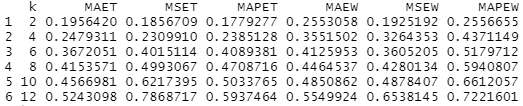
****

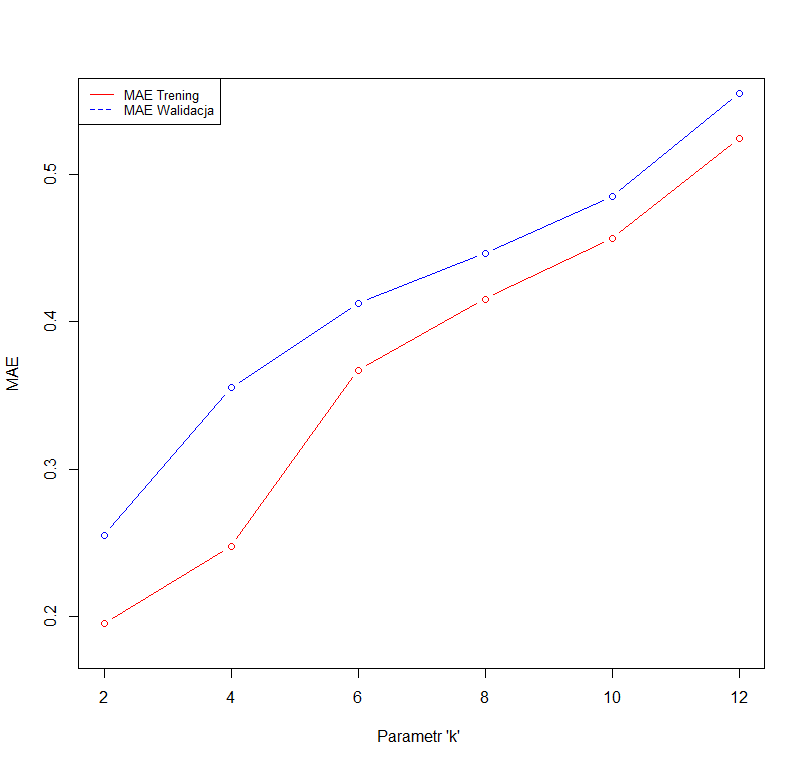
**PAKIET NNET – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**KNN kod własny – Regresja**

****

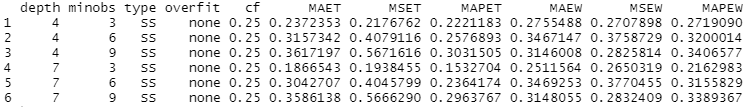
****

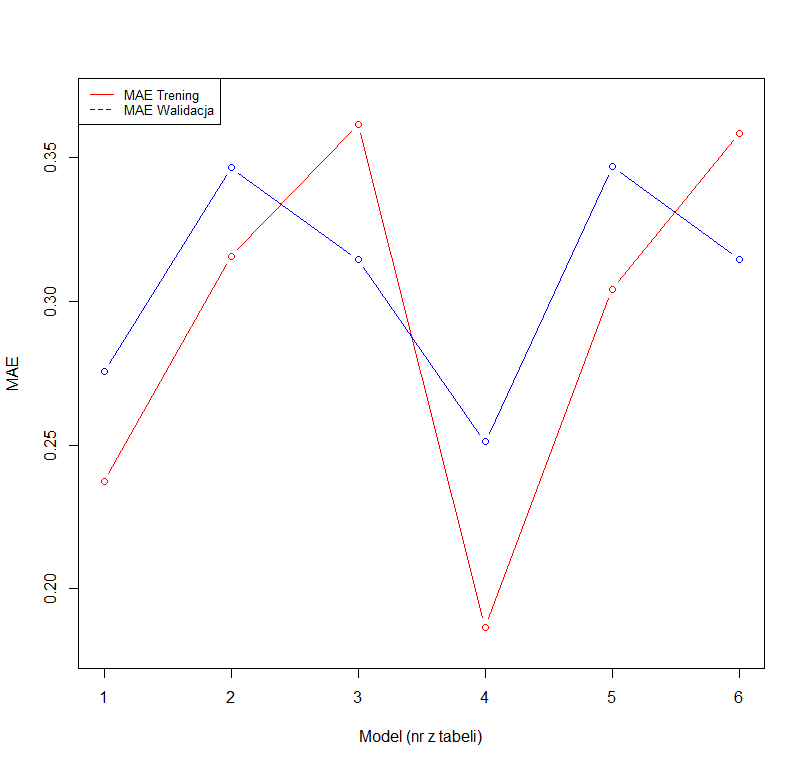
**PAKIET KNN – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**DRZEWA DECYZYJNE kod własny – Regresja**

****

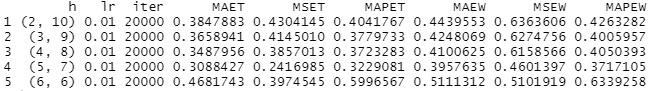
****

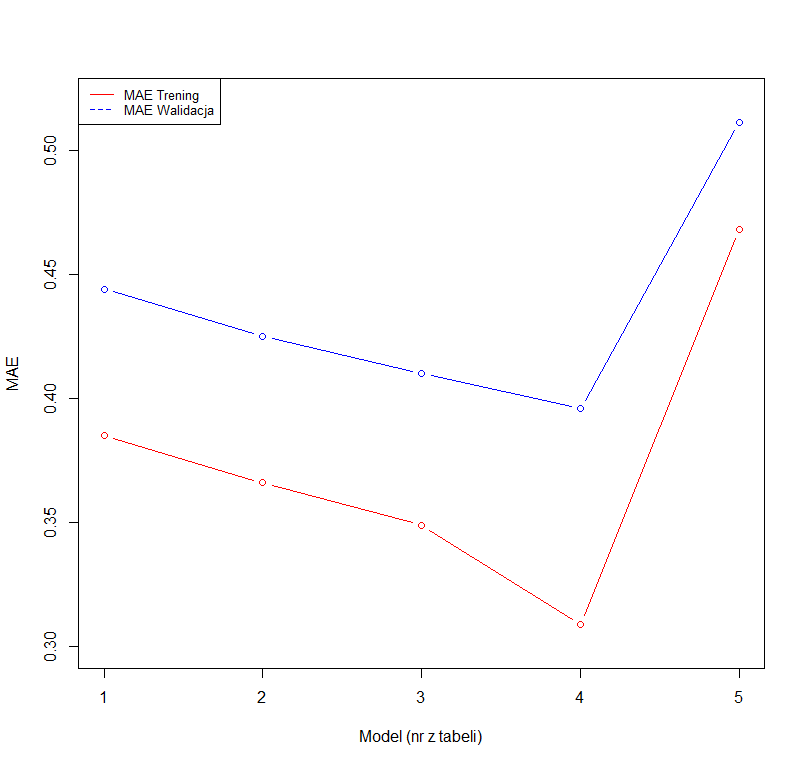
**PAKIET RPART – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**SIECI NEURONOWE kod własny – Regresja**

****

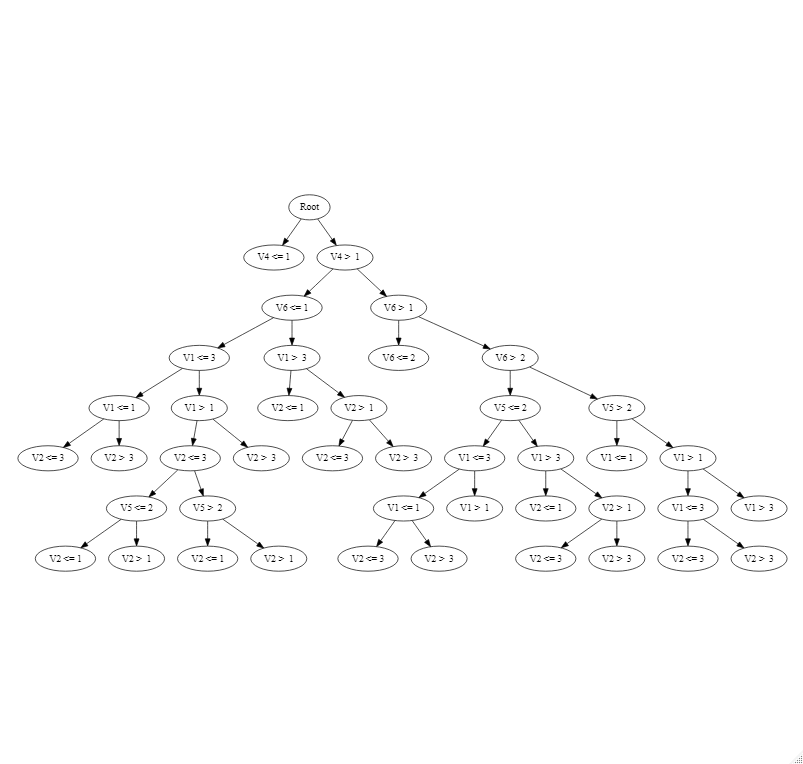
****

**PAKIET NNET – KROSWALIDACJA z pomocą funkcji ‘train’ – najlepszy model:**

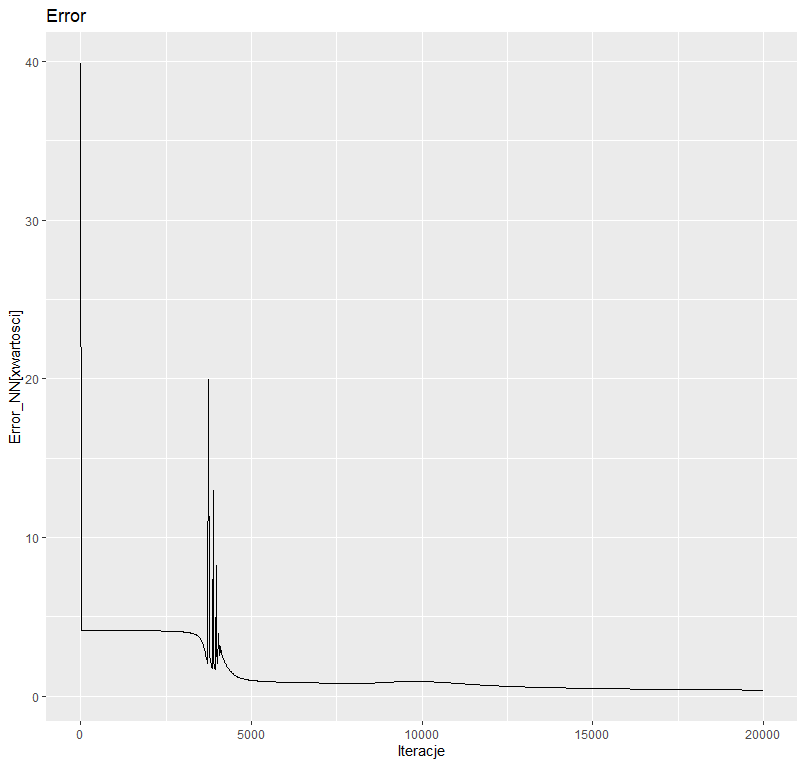
**Obraz zawierający tekst

Opis wygenerowany automatycznie**

**DRZEWA DECYZYJNE – Przykład drzewa (kod własny) na danych do klasyfikacji wieloklasowej   
 Tak wygląda drzewo (najlepszy model z własnej implementacji)**

****

**SIECI NEURONOWE – Przykład uczenia sieci (kod własny) na danych do regresji   
 Tak wygląda błąd podczas uczenia (najlepszy model z własnej implementacji)**

****