Autor: Maksymilian Romańczuk GD41582

Data: 04.11.19

Przedmiot: Sztuczna Inteligencja

Numer zadania: 1

Cel: Klasyfikacja poziomu nauki podanej w latach za pomocą sieci neuronowych, zadanie polega na sprawdzeniu, poziomu edukacji przy wynikach na bazie bezpłodności przy poronieniu bądź aborcji.

Opis:

Metoda oceniania: hold-out, lambda 2/3

Zbiór danych: datasets, infert

ToothGrowth: 8 kolumny, 248 wierszy

[,1] eductaion factor poziom edukacji(0-5/6-11/12+yrs)

[,2] age numeric wiek osoby badanej

[,3] parity numeric

[,4] inducted numeric wykonane aborcje

[,5] case numeric

[,6] spontaneous numeric poronienia

[,7] straturm numeirc

[,8] pooled.straturm numeric

Kod z komentarzem:

//instalacja pakietu amore

install.packeges("AMORE")

library(AMORE)

//generujemy dane uczące sieć

data("infert")

infert

table(infert$education)

l.danych=nrow(infert)

set.seed(8)

//40(2/3) losowych indexów do trenowania i pozostałe 20(1/3) do testowania

idxTren<-sample(1:l.danych,2\*l.danych/3)

idxTest<-setdiff(1:l.danych,idxTren)

//definiujemy funkcję zmieniającą zmienną zawierającą nazwę etykiety

//na liczbę zmiennych binarnych, ile jest wartości etykiet

target<-function(x)

{

n<-length(x)

wartosci<-levels(x)

l<-length(wartosci)

T<-matrix(0,nrow=n,ncol=l)

for(i in 1:l)

T[,i]<-(x==wartosci[i])

colnames(T)<-wartosci

return(T)

}

//zastosowanie funkcji dla danych określających etykiety

wZadane<-target(infert$education)

wZadane

set.seed(128)

//tworzymy strukturę sieci

siec<-newff(n.neurons=c(6,6,3),

learning.rate.global=0.03,

momentum.global=0.5,

hidden.layer="sigmoid",

output.layer="purelin",

method="ADAPTgdwm",

error.criterium="LMS")

//trenujemy sieć

wynik<-train(siec,

infert[idxTren,c(2,3,4,5,6,7,8)],

wZadane[idxTren,],

error.criterium="LMS",

report=TRUE,

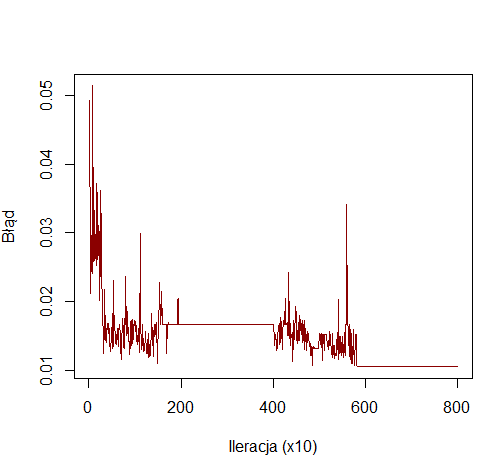
show.step=10,

n.shows=800)

//wyświetlam wartość błędów

plot(wynik$Merror,type="l",xlab="Ileracja (x10)",

ylab="Błąd", col="darkred")



//stosujemy wytrenowaną sieć do danych testowych

y<-sim(wynik$net,infert[idxTest, c(2,3,4,5,6,7,8)])

y

//definiuję funkcję oceny klasyfikacji (zamieniam liczby na etykietę)

test.klasyf<-function(zad,wy)

{

zadane<-max.col(zad)

rozpoznane<-max.col(wy)

print(table(zadane,rozpoznane))

}

wynik<-test.klasyf(wZadane[idxTest,],y)

//określam dokłądność klasyfikacji

cat("Dokładność klasyfikacji:",

sum(diag(wynik))/sum(wynik)\*100, "%\n")

Dokładność klasyfikacji: 97%

Podsumowanie: Klasyfikacja poziomu edukacji na poziomie 97% oznacza, że sieć dostała wystarczająca ilość danych i ”wyuczyła się” porównywać podane jej dane i oceniać w bardzo logiczny sposób, który daje prawie idealne wyniki, które przypuszcza użytkownik.