Autor: Maksymilian Romańczuk GD41582

Data: 04.11.19

Przedmiot: Sztuczna Inteligencja

Numer zadania: 1

Cel: Klasyfikacja wzrostu zębów świnki morskiej za pomocą sieci neuronowych, zadanie polega na sprawdzeniu, który sposób lepiej wpływa na wzrost zębów, przy użycie witaminy C czy soku pomarańczowego.

Opis:

Metoda oceniania: hold-out, lambda 2/3

Zbiór danych: datasets, ToothGrowth

ToothGrowth: 3 kolumny, 60 wierszy

[,1] len numeric długość zębów

[,2] supp factor typ suplementu (witamina C/sok pomarańczowy)

[,3] dose numeric dawka w miligramach na dzień

Kod z komentarzem:

//instalacja pakietu amore

install.packeges("AMORE")

library(AMORE)

//generujemy dane uczące sieć

data(ToothGrowth)

ToothGrowth

table(ToothGrowth$supp)

l.danych=nrow(ToothGrowth)

set.seed(112)

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

idxTren<-sample(1:l.danych,40)

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(ToothGrowth$supp)

wZadane

set.seed(122)

//tworzymy strukturę sieci

siec<-newff(n.neurons=c(2,2,2),

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,

ToothGrowth[idxTren,c(1,3)],

wZadane[idxTren,],

error.criterium="LMS",

report=TRUE,

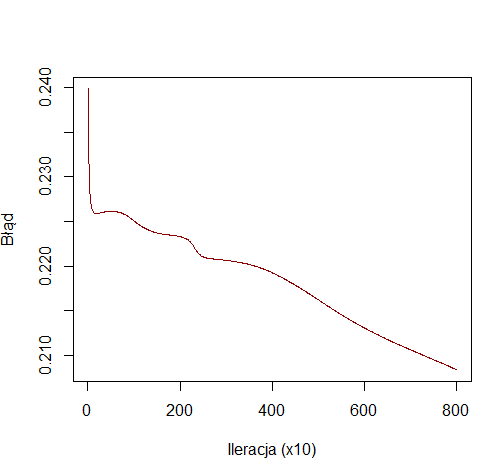
show.step=10,

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,ToothGrowth[idxTest, c(1,3)])

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: 65%

Podsumowanie: Klasyfikacja wzrostu zębów u świnek morskich na poziomie 65% oznacza, że sieć nie dostała wystarczającej ilości danych do klasyfikacji dającej racjonalne wyniki.