

EPC1: Perceptron

Guilherme Rocha Gonçalves¹, Rodolfo Coelho Dalapicola¹

Universidade de São Paulo
São Carlos, Brasil

1 Introdução

1.1 Contextualização

Primeira atividade do estudo sobre Redes Neurais Artificiais desenvolvida durante o curso da disciplina SEL5712 Redes Neurais Artificiais, oferecida no primeiro semestre do ano de 2019 para os alunos do programa de mestrado em Engenharia Elétrica e oferecido pelo Professor Dr. Ivan Nunes da Silva.

1.2 Objetivo

Esta atividade tem como objetivo a criação de uma rede neural Perceptron para solução de um problema de engenharia. Além disso, este trabalho pode ser considerado como um complemento ao que foi lecionado em sala de aula.

2 Materiais e Métodos

2.1 Problema

Uma rede neural artificial precisa ser treinada para automatizar a classificação de duas classes de pureza de um determinado óleo.

2.2 Bases de treino e de teste

O treinamento da rede neural será feito usando a *Regra de Aprendizado de Hebb*. A Regra de Hebb busca corrigir os pesos da rede neural quando a saída obtida difere da esperada. Esse ajuste é feito proporcionalmente ao valor das entradas. Esse tipo de método admite infinitas soluções, uma vez que qualquer variação nos pesos iniciais ou na taxa de aprendizagem pode gerar uma solução válida diferente. O Perceptron é usado para classificação de classes linearmente separáveis. Os dados de treino e de testes são fornecidos pelo EPC2.

2.3 Linguagem

Os códigos deste trabalho foram feitos usando *python*. Nenhuma biblioteca de Redes Neurais foram usadas, apenas bibliotecas para manipulação de matrizes e visualização de gráficos. O código está no Apêndice e também pode ser acessado através do link <https://github.com/grgoncal/SEL5712—Redes-Neurais-Artificiais>.

3 Resultados e discussão

3.1 Treinamento da Rede Perceptron

Foram feitos 5 treinamentos para a rede. O vetor de pesos foi iniciado com valores aleatórios entre 0 e 1, a taxa de aprendizagem foi definida como 0.01. O resultado dos treinamentos estão documentado na Tabela 1.

Treino	w0	w1	w2	w3	w0	w1	w2	w3	Épocas
1	0.13230535	0.21819187	0.99455965	0.18421044	-3.04769465	1.55439387	2.46076965	-0.72791756	377
2	0.13824052	0.3883754	0.87875059	0.73382287	-2.92175948	1.4098594	2.41328459	-0.69933113	349
3	0.0764187	0.97292242	0.63733467	0.04418705	-3.1035813	1.55359642	2.48919267	-0.73880895	390
4	0.97040838	0.64609153	0.50880691	0.9028419	-3.16959162	1.60609953	2.55031291	-0.7538541	471
5	0.24839085	0.08900118	0.97372129	0.39714327	-2.93160915	1.41518518	2.43127329	-0.70201073	368

Tabela 1: Treinamento

3.2 Perceptron em casos de testes

Foram fornecidos alguns objetos de teste para que a rede possa classificar após cada iteração de treinamento. Os resultados foram iguais para todos os casos, como era esperado. A Tabela 2 mostra os resultados.

Amostra	x1	x2	x3	y(1)	y(2)	y(3)	y(4)	y(5)
1	-0.3565	0.062	5.9891	-1	-1	-1	-1	-1
2	-0.7842	1.1267	5.5912	1	1	1	1	1
3	0.3012	0.5611	5.8234	1	1	1	1	1
4	0.7757	1.0648	8.0677	1	1	1	1	1
5	0.157	0.8028	6.304	1	1	1	1	1
6	-0.7014	1.0316	3.6005	1	1	1	1	1
7	0.3748	0.1536	6.1537	-1	-1	-1	-1	-1
8	-0.692	0.9404	4.4058	1	1	1	1	1
9	-1.397	0.7141	4.9263	-1	-1	-1	-1	-1
10	-1.8842	-0.2805	1.2548	-1	-1	-1	-1	-1

Tabela 2: Resultado dos testes.

3.3 Diferença de épocas no Perceptron

Diferente do ADALINE, o perceptron permite que existam infinitas soluções. Portanto, dependendo dos parâmetros estáticos, tais como a taxa de aprendizagem e os valores iniciais do vetor de pesos, o número de épocas pode variar.

Se a taxa de aprendizagem for muito baixa, pode ser que mais épocas sejam necessárias para uma solução ser encontrada. Caso a taxa de aprendizagem for muito alta, o algoritmo pode ficar instável, aumentando em alguns casos o número de épocas ou na pior das hipóteses, impossibilitando a convergência. No algoritmo desenvolvido para esse trabalho, o número de épocas aumentou quando a taxa de aprendizagem foi alterada para 0.001, chegando em até 8 épocas, em alguns casos.

4 Conclusão

A grande desvantagem do perceptron é exigir que as classes a serem classificadas sejam obrigatoriamente linearmente separáveis. Caso contrário, a rede não converge.

Além dessa limitação, foi verificado pelo proponente a influência da taxa de aprendizagem na convergência e também a importância de se limitar o número de épocas, para eventuais erros no código ou mesmo para evitar *loops* infinitos causados pela não-convergência.

O Perceptron é um ótimo início de rede para a disciplina e apesar de simples, pode resolver muitos problemas.

5 Apêndice

5.1 Código fonte

O código fonte dos algoritmos desenvolvidos está disponível no repositório público [1] para livre acesso.

```

1 # IMPORTS -----
2 import pandas as pd
3 import numpy as np
4
5 #////////////////////////////////////
6 # CONSTANTS //////////////////////////////////
7 #////////////////////////////////////
8
9 learningRate = 100
10 maxEpochs = 1000
11 errTol = 0
12 trainNumber = 5
13
14 #////////////////////////////////////
15 # TRAIN //////////////////////////////////
16 #////////////////////////////////////
17
18 def train(x, d, w):
19     epoch = 0
20     err = 1
21     print ("\n[W INITIAL]   " + str(w.T.values))
22
23     while(epoch < maxEpochs and err > errTol):
24         err = 0
25         for i in range(0,x.shape[0]):
26             u = w.iloc[:,0] * x.iloc[i,:]
27             u = u.sum()
28             if u >= 0:
29                 y = 1
30             else:
31                 y = -1
32
33             if abs(y - d.iloc[i]) == 2:
34                 w = (w.T + learningRate * (d.iloc[i] - y) * x.
35                     iloc[i,:]).T
36                 err += 1
37
38             # print err
39             epoch += 1
40
41         print (" [W FINAL]   " + str(w.T.values))
42         print (" [EPOCHS]   " + str(epoch) + "\n")
43         return w
44
45 #////////////////////////////////////
46 # TEST //////////////////////////////////
47 #////////////////////////////////////
48
49 def test(w):
50     x = pd.read_csv('./test.csv', header = None)

```

```

50     for name in reversed(x.columns.values):
51         x.rename(columns = {name : name + 1}, inplace = True)
52     x.insert(loc = 0, column=0, value = np.full((x.shape[0],1)
53         , -1))
54
55     y = []
56
57     for i in range(0, x.shape[0]):
58         u = w.iloc[:,0] * x.iloc[i,:]
59         u = u.sum()
60         if u >= 0:
61             y.append(1)
62         else:
63             y.append(-1)
64
65     return y
66
67 # MAIN
68
69
70 # TRAIN DATASET
71 data = pd.read_csv('./train.csv', header = None)
72
73 # GET INPUTS, OUTPUTS AND WEIGHTS
74 x = data.iloc[:,1:(data.shape[1] - 1)]
75     # GET INPUTS
76     x.insert(loc = 0, column=0, value = np.full((x.shape[0],1),
77         -1))
78     # ADD -1 INPUT
79     d = data.iloc[:,(data.shape[1] - 1)]
80     # GET OUTPUTS
81
82     for i in range(1, trainNumber + 1):
83         w = pd.DataFrame(np.random.rand((data.shape[1] - 1)))
84         # GENERATE WEIGHTS
85         print("[TRAINING NUMBER " + str(i) + "]")
86         y = test(train(x, d, w))
87         print("[RESULT OF TRAINING NUMBER " + str(i) + "] " + str(
88             y) + "\n")

```

Listing 1.1: Python code

Referências Bibliográficas

- [1] Gonçalves, G.R.: Github: Sel5712 - redes neurais artificiais. github.com/grgoncal/SEL5712—Redes-Neurais-Artificiais, acessado: 2019-04-05