

Relatório de Classificação de Padrões

Victor Carreira

Resumo—Este relatório apresenta o resultado de sete testes conduzidos em uma rede neural artificial *Perceptron multi-camadas* em um banco de dados de uma instituição financeira com o intuito de se fazer uma análise de crédito bancário. A base de dados contém 2077 exemplos de créditos concedidos. Possui 11 atributos de entrada e 2 classes de saída. A saída da rede indica se o cliente pagou o empréstimo ou não. Os resultados dos testes estão apresentados em formato de tabela.

I. INTRODUÇÃO

As Redes Neurais Artificiais (RNA) são inspiradas em modelos sensoriais do processamento de tarefas realizadas pelo cérebro (Hagan et al., 1996). Uma RNA, portanto pode ser criada através da aplicação de algoritmos matemáticos que imitem a tarefa realizada por um neurônio (Nedjah et al., 2016). Uma rede neural artificial possui semelhanças com a rede biológica presente no sistema nervoso central, neste o cômputo de informações realizado do cérebro é feito através de uma vasta quantidade de neurônios interconectados (Feldman et al., 1988; Poulton, 2002). A comunicação entre essas células é realizada através de impulsos elétricos. Estes são transmitidos e recebidos por meio de sinapses nervosas entre axônios e dendritos. As sinapses são estruturas elementares e uma unidade funcional localizada entre dois neurônios (Krogh, 2008).

McCulloch and Pitts (1943) redigem o trabalho pioneiro onde foi modelado um neurônio cuja resposta dependia do *input*¹ que provinha de outros neurônios e do peso utilizado. Já Rosenblatt (1962) cria a teoria de convergência do *Perceptron* onde ele prova que modelos de neurônios possuem propriedades similares ao cérebro humano (Kanal, 2001). Neste sentido as rede neuronais artificiais podem realizar performances sofisticadas no reconhecimento de padrões, mesmo se alguns neurônios forem destruídos (Levy, 1997). Minsky and Papert (1969) demonstraram que *Perceptrons* somente resolvem uma classe muito limitada de problemas que podem ser linearizados.

II. OBJETIVO E METODOLOGIA

Uma instituição financeira possui uma base de dados com o histórico de crediário oferecido aos seus clientes. Baseado neste histórico, a instituição deseja inferir se um novo cliente pagará ou não a dívida contraída.

A base de dados possui 2077 exemplos, com 11 atributos cada, de créditos concedidos aos seus clientes. A base informa ainda se o cliente honrou ou não o pagamento do empréstimo.

A partir da base original, foram criadas 3 bases de treinamento, com 1500 exemplos cada escolhidos aleatoriamente a

partir da base original, e 3 bases de testes com 577 exemplos cada, representando, respectivamente, 72, 2% e 27, 8% do total de cada sub- grupos de dados. Estas bases estão nos arquivos treino01.txt, treino02.txt, treino03.txt, teste01.txt, teste02.txt e teste03.txt. Utilizar o software WEKA, para criar um classificador, baseado em redes neurais, capaz de informar se um novo cliente será potencialmente adimplente ou não.

Abra cada um dos arquivos no WEKA e grave-os em formato .arff. Através de um editor de textos (por exemplo, o WordPad), altere o tipo associado às variáveis categóricas conforme o exemplo `@attribute NDEP numeric => @attribute NDEP 0,1,2,3,4,5,6,7`

Para cada uma das configurações abaixo, apresente os resultados para cada par de conjuntos de treino e de teste, assim como a média e o desvio padrão dos 3 pares: I, II, III, IV, V, VI.

- Sem normalização dos atributos de entrada;
- Com normalização dos atributos de entrada e SEM codificação binária dos atributos categóricos;
- Com normalização e codificação binária dos atributos categóricos de entrada e com 2 números diferentes de neurônios na camada escondida;
- Com normalização dos atributos de entrada e variando o número de épocas durante a fase de treinamento. Escolha 3 durações de treino diferentes (por exemplo: 1, 100 e 1000);
- Com normalização dos atributos de entrada e utilizando um conjunto de validação;
- Tente obter melhores resultados (se possível) agrupando algumas categorias das variáveis ESTC e NDEP. Para isto, utilize o filtro não supervisionado MergeTwoValues.

Para os itens I, II, IV, V e VI, indique para cada um dos casos o número de neurônio na camada escondida e explique a sua escolha. Para todos os itens, não varie a taxa de aprendizagem nem o termo de momento.

III. PRINCÍPIO TEÓRICO

O neurônio de McCulloch and Pitts (1943) propõe um limite binário para a criação de um modelo. Este neurônio artificial registra uma soma de pesos de n sinais de entrada, x_j , $j = 1, 2, 3, \dots, n$, e fornece um *output*² de 1 caso esta soma esteja acima do limite u . Caso contrário o *output* é 0. Matematicamente essa relação pode ser descrita de acordo com a Eq. 1:

$$y = \theta \left(\sum_{j=1}^n w_j x_j - u \right) \quad (1)$$

Professora: Marley. Eng. Elétrica. PUC-RIO

¹Valor de entrada

²Valor de saída

Onde θ é o passo dado na posição 0, w_j é chamada sinapse-peso associado a um *jésimo* input. A título de simplificação a função limite³ u é considerada um outro peso $w_0 = -u$ anexado a um neurônio com um input constante $x_0 = 1$. Pesos positivos correspondem a uma sinapse **excitatória**, enquanto pesos negativos correspondem a uma sinapse **inibitória**. Este modelo contém uma série de simplificações que não refletem o verdadeiro comportamento dos neurônios biológicos (Mao, 1996).

Derivações do neurônio de McCulloch and Pitts (1943) na escolha das funções de ativação. Uma função largamente utilizada é a função sigmóide, que exibe uma suavização dos *outputs* a medida que o valor da função diminui (Mao, 1996; Misra and Saha, 2010). Essa função de ativação pode ser expressa de acordo com a Eq. 2:

$$g(x) = 1/(1 + e^{-\beta x}) \quad (2)$$

Onde β é o parâmetro de inclinação. A Fig. 1 ilustra a sequência lógica da operação de uma RNA para um neurônio simples de McCulloch-Pitts.

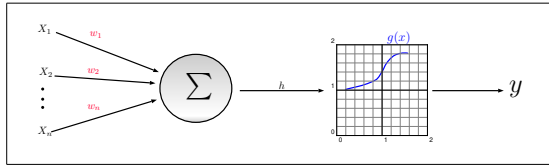


Figura 1. Modelo esquemático de um neurônio de McCulloch-Pitts. Onde x_1, x_2, \dots, x_n são os *inputs*, w_1, w_2, \dots, w_n são os pesos, h é o treino, $g(x)$ é a função de ativação, e y é o *output*.

As redes alimentadas diretamente são aquelas redes cujos grafos orientados se distinguem pela presença de um ou mais camadas ocultas e cujos nós são chamados de neurônios ocultos. A função do neurônio é intervir entre a camada externa e a saída da rede de maneira útil. Adicionando-se camadas ocultas a rede torna-se capaz de realizar estatísticas de ordem elevada (Haykin, 1999).

IV. RESULTADOS

A tabela I refere-se ao conjunto de 6 testes realizados no primeiro banco de dados que é uma composição dos dados de treinamento e testes 1. Os testes utilizados neste trabalho foram em sua totalidade compostos por 2 camadas ocultas. O número de neurônios ocultos estão indicados nas tabelas, como Neurônios na CO1 (camada oculta 1) e como Neurônios na CO2 (camada oculta 2).

³Genericamente chamada de função de ativação

Tabela I
GRUPO 01

Configuração	I	II	III	IV	V	VI
Neurônios na CO1	4	4	4	-	4	4
Neurônios na CO2	4	4	2	-	4	4
Cl. Correta (%)	52.3	91.3	91.2	-	52.3	97
Cl. Incorreta (%)	47.7	8.7	8.8	-	47.7	3
MAE (%)	0.49	0.13	0.14	-	0.49	0.04
RMSE (%)	0.50	0.26	0.27	-	0.49	0.17
RAE (%)	99.68	26.23	28.3	-	100.2	8.50
RRSE(%)	100.23	52.23	53.75	-	100.1	36.51
Dados Totais	1500	1500	1500	-	1500	100

A tabela II refere-se os testes de variações de épocas do conjunto de dados 1.

Tabela II
GRUPO 01: VARIANDO AS ÉPOCAS NO TESTE IV

Configuração	IV ₁	IV ₁₀₀	IV ₁₀₀₀
Neurônios na CO1	4	4	4
Neurônios na CO2	4	4	4
Cl. Correta (%)	52.3	90.8	91.53
Cl. Incorreta (%)	47.7	9.2	8.47
MAE (%)	0.49	0.16	0.14
RMSE (%)	0.51	0.27	0.26
RAE (%)	99.63	32.34	27.49
RRSE(%)	100.23	54.49	52.89
Dados Totais	1500	1500	1500

Aqui iniciam-se o conjunto de testes 2.

Tabela III
GRUPO 02

Configuração	I	II	III	IV	V	VI
Neurônios na CO1	4	4	4	-	4	4
Neurônios na CO2	4	4	8	-	4	4
Cl. Correta (%)	53.2667	91.4667	91.6667	-	53.2667	91.8667
Cl. Incorreta (%)	46.7333	8.5333	8.3333	-	46.7333	8.1333
MAE (%)	0.4927	0.1247	0.1213	-	0.4999	0.1165
RMSE (%)	0.5052	0.2637	0.2657	-	0.4999	0.2584
RAE (%)	98.9582	25.0564	24.3633	-	100.4037	23.3909
RRSE(%)	101.2566	52.8491	53.2471	-	100.1905	51.7846
Dados Totais	1500	1500	1500	-	1500	1500

Variações das épocas do conjunto de testes 2.

Tabela IV
GRUPO 02: VARIANDO AS ÉPOCAS NO TESTE IV

Configuração	IV ₁	IV ₁₀₀	IV ₁₀₀₀
Neurônios na CO1	4	4	4
Neurônios na CO2	4	4	4
Cl. Correta (%)	53.2667	91.4	91.8667
Cl. Incorreta (%)	46.7333	8.6	8.1333
MAE (%)	0.4913	0.1384	0.1165
RMSE (%)	0.5083	0.2669	0.2584
RAE (%)	98.6867	27.807	23.3909
RRSE(%)	101.8876	53.4966	51.7846
Dados Totais	1500	1500	1500

Aqui iniciam-se o conjunto de testes 3.

Tabela V
GRUPO 03

Configuração	I	II	III	IV	V	VI
Neuronios na CO1	4	4	1	-	4	4
Neuronios na CO2	4	4	12	-	4	4
Cl. Correta (%)	49	91.4	90.8667	-	51	66
Cl. Incorreta (%)	51	8.6	9.1333	-	49	34
MAE (%)	0.5006	0.1228	0.1638	-	0.5	0.4979
RMSE (%)	0.5013	0.2654	0.2783	-	0.5	0.4979
RAE (%)	100.151	24.5646	32.7792	-	100.0321	110.5297
RRSE(%)	100.2846	53.0866	55.6619	-	100.0128	104.9763
Dados Totais	1500	1500	1500	-	1500	100

Rosenblatt, F. (1962). Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms. *Archives of General Psychiatry*, 7:218–219.

Aqui seguem os testes de variação de épocas.

Tabela VI
GRUPO 03: VARIANDO AS ÉPOCAS NO TESTE IV

Configuração	IV ₁	IV ₁₀₀	IV ₁₀₀₀
Neuronios na CO1	4	4	4
Neuronios na CO2	4	4	4
Cl. Correta (%)	49	90.3333	91
Cl. Incorreta (%)	51	9.6667	9
MAE (%)	0.5002	0.1321	0.1318
RMSE (%)	0.5004	0.2748	0.2653
RAE (%)	100.0891	26.4304	26.3714
RRSE(%)	100.099	54.9758	53.0768
Dados Totais	1500	1500	1500

V. CONCLUSÕES

Ao analisar previamente os dados é possível que o número de dependentes NDEP é majoritariamente composto por uma pessoa

REFERÊNCIAS

- Feldman, J. A., Fanty, M. A., and Goddard, N. H. (1988). Computing With Structured Neural Networks. *Computer*, 21(3):91–103.
- Hagan, M. T., Demuth, H. B., Beale, M. H., and De Jesús, O. (1996). Neural Network Design. page 1012.
- Haykin, S. (1999). Prentice Hall Inc., Porto Alegre, 2 edition.
- Kanal, L. N. (2001). Perceptrons. *Encyclopedia of Computer Science*, pages 11215–11218.
- Krogh, A. (2008). What are artificial neural networks? *Nature biotechnology*, 26(2):195–197.
- Levy, S. (1997). The Computer. *Newsweek*, 130(22):28.
- Mao, J. (1996). Why artificial neural networks? *Communications*, 29:31–44.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133.
- Minsky, M. and Papert, S. (1969). *Perceptrons*. 2 edition.
- Misra, J. and Saha, I. (2010). Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing*, 74(1-3):239–255.
- Nedjah, N., da Silva, F. P., de Sá, A. O., Mourelle, L. M., and Bonilla, D. A. (2016). A massively parallel pipelined reconfigurable design for M-PLN based neural networks for efficient image classification. *Neurocomputing*, 183:39–55.
- Poulton, M. M. (2002). Neural networks as an intelligence amplification tool: A review of applications. *Geophysics*, 67(3):979.