

Rede Neural Perceptron Simples

Savio Lopes Rabelo

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Campus Fortaleza – CE – Brasil

savio.rabelo@ppgcc.ifce.edu.br

Resumo. Este relatório descreve a implementação da rede neural Perceptron Simples aplicada à classificação de padrões. A metodologia utilizada para a implementação é constituída por duas fases: treinamento e teste, com cada conjunto sendo composto por 80% e 20% das bases de dados, respectivamente. Foram usadas quatro bases de dados disponíveis online no repositório UCI Machine Learning. Os resultados são bastante satisfatórios, chegando em taxas de acerto em 100% em algumas bases.

1. Introdução

Basicamente o Perceptron consiste de uma única camada de neurônios com pesos sinápticos e bias ajustáveis. Se o padrão de entrada forem linearmente separáveis, o algoritmo de treinamento possui convergência garantida, isto é, tem capacidade para encontrar um conjunto de pesos que classifica corretamente os dados. Os neurônios do perceptron são similares ao de McCulloch-Pitts, por terem a função de ativação do tipo degrau, mas possuem pesos associados e bias.

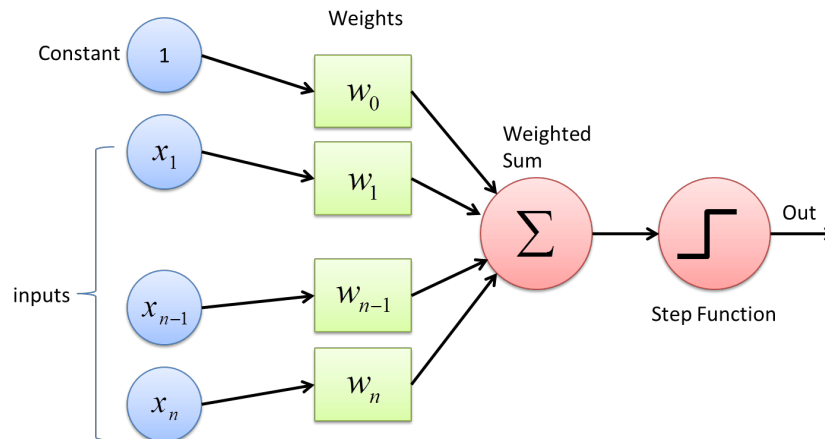


Figura 1. Rede Neural Perceptron Simples [Towardsdatascience 2019].

Para cada padrão de treinamento x_i , a saída da rede y_i é calculada. Determina-se o erro e_i entre a saída desejada para esse padrão d_i e a saída da rede y_i , $e_i = d_i - y_i$. O vetor de pesos é atualizado de acordo com a seguinte regra:

$$\vec{w}_i(t+1) = \vec{w}_i(t) + \eta e_i \vec{x}_i. \quad (1)$$

2. Metodologia

No primeiro momento foi realizada a separação do conjunto de dados em dois subconjuntos: treinamento e teste. Os valores utilizados para os conjuntos equivalem a 80% do conjunto original para a fase de treinamento e 20% do conjunto original para a fase de teste. Logo depois, os dados foram normalizados para eliminação de redundâncias indesejadas e também foram embaralhados.

Para a avaliação dos resultados alcançados na classificação, foram utilizados as seguintes métricas: a precisão ou valor preditivo positivo, taxa de sensibilidade ou taxa positiva verdadeira, especificidade ou taxa real negativa e acurácia. As Equações são apresentadas a seguir:

$$Precisao = \frac{VP}{VP + FP}, \quad (2)$$

$$Sensibilidade = \frac{VP}{VP + VN}, \quad (3)$$

$$Especificidade = \frac{VN}{N} = \frac{VN}{FP + VN}, \quad (4)$$

$$Acuracia = \frac{VP + VN}{P + N}, \quad (5)$$

onde P e N é o número de padrões de cada classe. VP é o verdadeiro positivo. VN é o verdadeiro negativo. FP é o falso positivo e FN é o falso negativo.

3. Conjuntos de Dados

Para análise comparativas neste estudo, foram usados quatro conjuntos de dados: *Iris Flower Data Set*, *Vertebral Column Data Set*, *Dermatology Data Set* e *Breast Cancer Wisconsin Data Set*; todos disponíveis online no repositório *UCI Machine Learning* [Lichman 2013].

O banco de dados da Íris¹ é o conjunto mais conhecido que se encontra na literatura de reconhecimento de padrões. O conjunto de dados contém 3 classes de 50 instâncias cada, onde cada classe se refere a um tipo de planta de íris. Uma classe é linearmente separável das outras 2 classes.

Informações dos atributos:

1. Tamanho da sépala em cm
2. Largura da sépala em cm
3. Tamanho da pétala em cm
4. Largura da pétala em cm
5. Classe:
 - (a) Iris Setosa
 - (b) Iris Versicolour

¹Disponível em <https://archive.ics.uci.edu/ml/datasets/iris>

(c) Iris Virginica

Já o conjunto de dados da Coluna Vertebral² contém seis valores para características biomecânicas usadas para classificar pacientes ortopedistas em 3 classes (normal, hérnia de disco ou espondilolistese) ou 2 classes (normal ou anormal). Foi utilizado nessa prática o conjunto com três classes.

Informações dos atributos:

1. Incidência pélvica
2. Inclinação pélvica
3. Ângulo de lordose lombar
4. Inclinação sacra
5. Raio pélvico
6. Grau de espondilolistese
7. Classe:
 - (a) Hérnia de Disco (DH)
 - (b) Espondilolistese (SL)
 - (c) Normal (NO)
 - (d) Anormal (AB)

O banco de dados de Dermatologia³ é constituído de 34 atributos. Esse banco é parte de um estudo que aponta o tipo de Eryhemato-Squamous Disease, uma doença de pele.

Informações dos atributos (valores de 0 a 3, exceto quando indicado):

1. Eritema
2. Escala
3. Fronteiras Definidas
4. Coceira
5. Fenômeno Koebner
6. Pápulas Poligonais
7. Pápulas Foliculares
8. Envolvimento da Mucosa Oral
9. Envolvimento no Joelho e no Cotovelo
10. Envolvimento do Couro Cabeludo
11. Histórico Familiar (0 ou 1)
12. Atributos Histopatológicos
- ⋮
33. Atributos Histopatológicos
34. Idade (Classe de 1 a 6)

E o banco de dados de Câncer de Mama⁴ é constituído de 10 atributos. Informações dos atributos:

1. Número do código de amostra (número de identificação)

²Disponível em <http://archive.ics.uci.edu/ml/datasets/vertebral+column>

³Disponível em <http://archive.ics.uci.edu/ml/datasets/dermatology>

⁴Disponível em <https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\%28Original\%29>

2. Clump Espessura (1 - 10)
3. Uniformidade do tamanho da célula (1 - 10)
4. Uniformidade da forma da Célula (1 - 10)
5. Adesão Marginal (1 a 10)
6. Tamanho Único de Células Epiteliais (1 - 10)
7. Núcleos Nus (1 - 10)
8. Cromatina Branda (1 a 10)
9. Nucleoli Normal (1 - 10)
10. Mitoses (1 - 10)
11. Classe (2 para benigno, 4 para maligno)

4. Simulações Computacionais

Para realizar os experimentos, foi utilizado um computador com a seguinte configuração: processador Intel(R) Core(TM) i7-6500U a 2.5 GHz com 8 GB de RAM e executando Windows 10. Além disso, foi utilizado a linguagem de programação Python. Todos os testes foram feitos com 50 realizações em cada base, foram realizadas no máximo 200 épocas, e o parâmetro η (taxa de aprendizado) foi atualizado de acordo com a Equação 6.

$$\eta = \eta_f \left(\frac{\eta_i}{\eta_f} \right)^{\left(\frac{e}{E} \right)} = 0.5 \left(\frac{0.05}{0.5} \right)^{\left(\frac{e}{200} \right)}. \quad (6)$$

Para análise comparativas neste estudo, foram usados quatro conjuntos de dados: *Iris Flower Data Set* em sua versão binária (setosa vs outras), *Vertebral Column Data Set* em sua versão binária (paciente patológico vs paciente normal), *Dermatology Data Set* (paciente patológico vs paciente normal) e *Breast Cancer Wisconsin Data Set* (paciente com câncer benigno vs paciente com câncer maligno).

A Tabela 1 mostra os resultados do classificador Perceptron Simples em todas as bases de dados, levando em consideração as métricas já mencionadas.

Métricas	Bases de Dados				
	Íris	Coluna	Dermatologia	Câncer	Artificial
Acurácia	99,60	85,29	99,73	96,00	100,00
Taxa Mínima	96,67	75,68	95,59	91,23	100,00
Taxa Máxima	100,00	93,32	100,00	99,12	100,00
Desvio Padrão	00,01	00,06	00,01	00,02	00,00
Sensibilidade	99,01	88,43	99,26	95,15	100,00
Especificidade	99,80	72,50	99,95	96,15	100,00
Precisão	99,63	87,24	99,68	94,42	100,00

Tabela 1. Resultados do classificador Perceptron Simples.

Na Figura 2 é apresentada a superfície de decisão com o classificador Perceptron Simples.

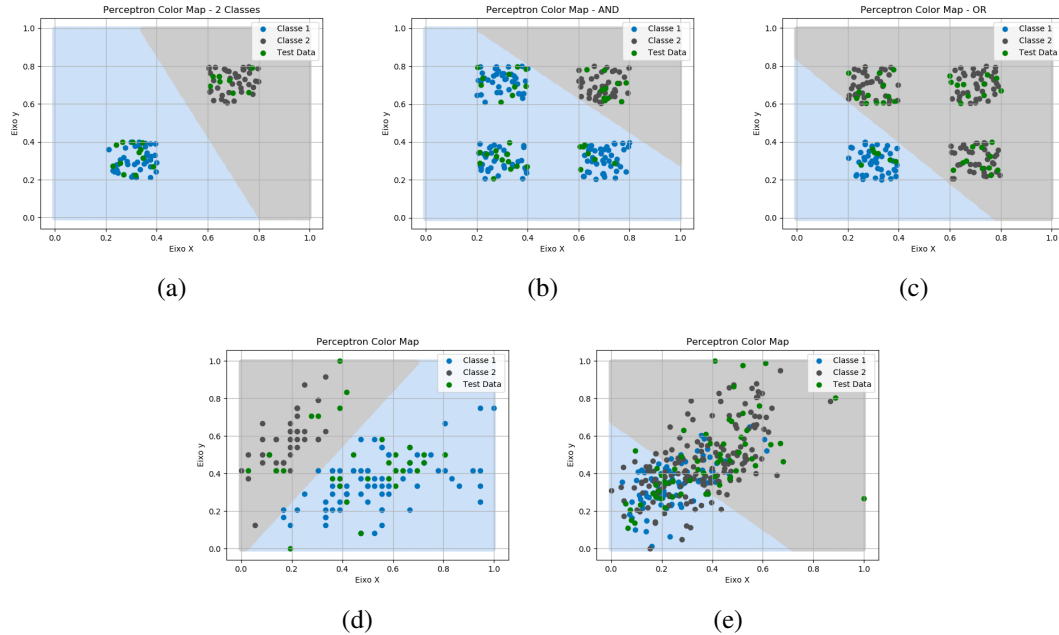


Figura 2. Superfície de decisão do classificador Perceptron Simples. (a) Base artificial com 2 classes. (b) Base artificial AND. (c) Base artificial OR. (d) Íris com dois primeiros atributos. (e) Coluna com dois primeiros atributos.

Na Figura 3 é apresentada a matriz de confusão com o classificador Perceptron Simples.

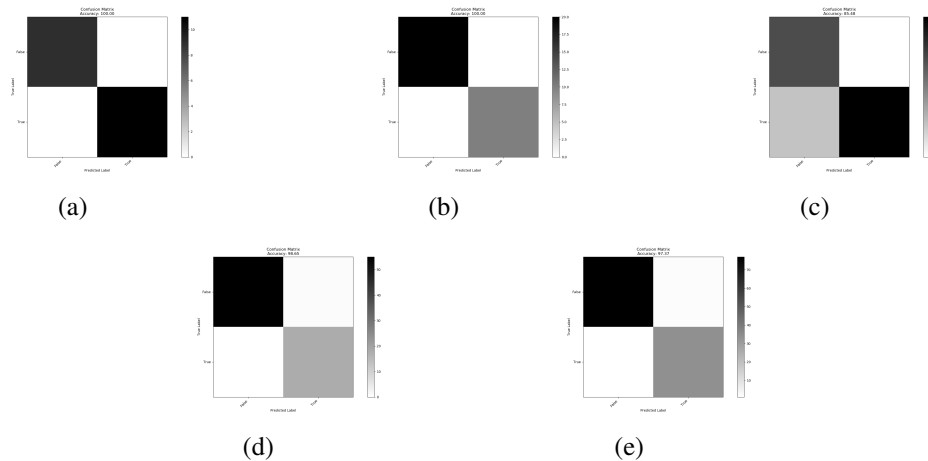


Figura 3. Matriz de confusão do classificador Perceptron Simples. (a) Base artificial com 2 classes. (b) Íris. (c) Coluna. (d) Dermatologia. (e) Câncer.

Na Figura 4 é apresentado um gráfico com erros na fase de treinamento do classificador Perceptron Simples.

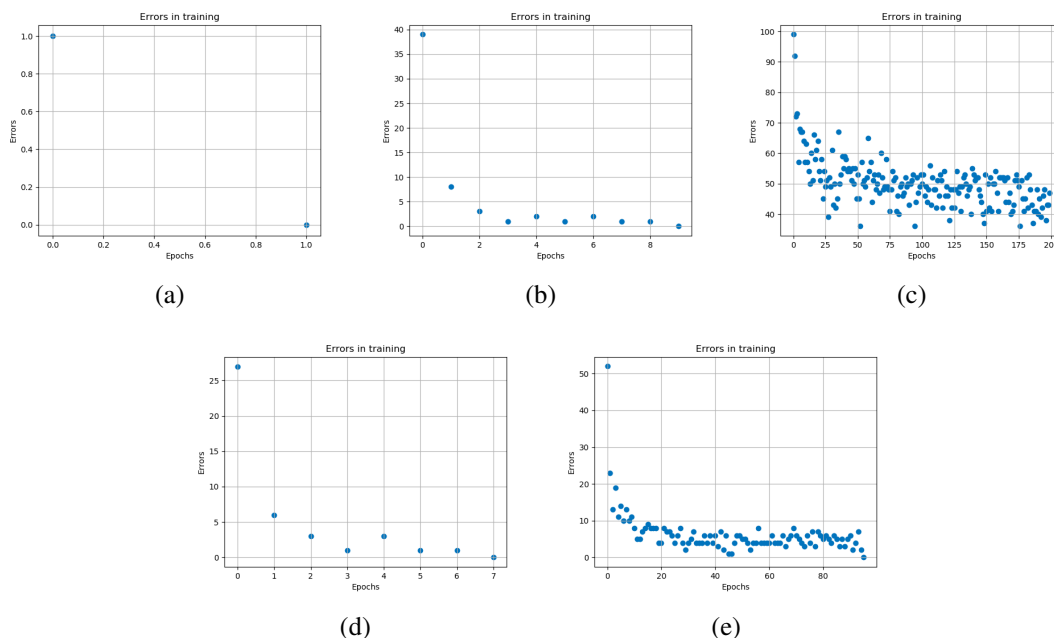


Figura 4. Erros na fase de treinamento do classificador Perceptron Simples. (a) Base artificial com 2 classes. (b) Íris. (c) Coluna. (d) Dermatologia. (e) Câncer.

5. Resultados

Analisando os experimentos, pode-se visualizar que em geral, o Perceptron Simples se sai muito bem na classificação de padrões, chegando a taxas de acerto de 100% quase em sua totalidade na base da Íris, como mostra na Tabela 1. Na base da Dermatologia e do Câncer foi obtido ótimo desempenho. Já com a base de dados gerada artificialmente, o resultado já era esperado obter 100% de acerto, visto que as classes são separáveis uma das outras.

Um dos motivos disso dá-se pela disposição dos dados. A base Iris tem as classes bastante separadas, tornando fácil encontrar uma reta que separe-as. Já em bases como a Coluna Vertebral, onde dados de diferentes classes sobrepõem-se, pode-se ver que o desempenho da taxa de acerto decai, ficando às vezes em torno de 78%-85%, com um alto desvio padrão, o que mostra sua instabilidade. Dessa forma, recomenda-se o uso da rede Perceptron simples, para classificação de padrões que não sobreponham-se.

Referências

- Lichman, M. (2013). UCI machine learning repository. <http://archive.ics.uci.edu/ml>.
- Towardsdatascience (2019). What the hell is perceptron. Disponível em <https://towardsdatascience.com/what-the-hell-is-perceptron-626217814f53>. Acesso em setembro de 2019.