
Aprendizagem de Máquina

Programa de Pós-Graduação em Engenharia Elétrica e de Computação (PPGEEC)

Introdução à Ciências de Dados - UFC *Campus* Sobral – 2023.1

Andressa Gomes Moreira – andressagomes@alu.ufc.br



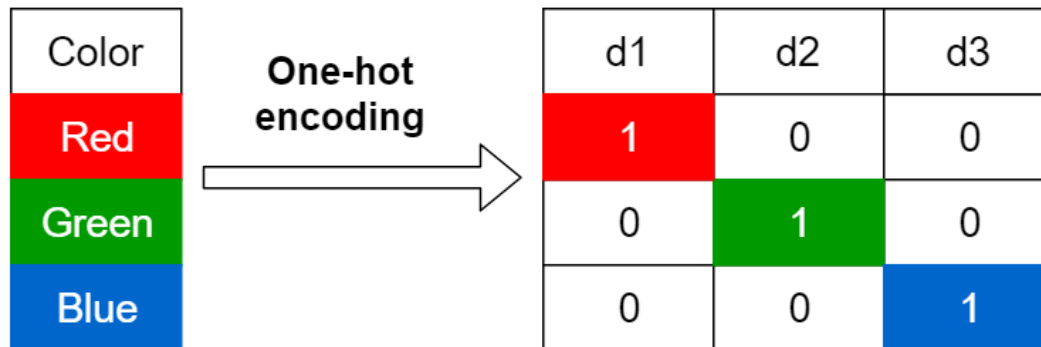
Sumário

1. Pré-processamento dos Dados
2. Algoritmos de Classificação
3. Avaliação dos Classificadores
4. Referências

Pré-processamento dos Dados

○ One-Hot Encoding

- A codificação one-hot é uma etapa de pré-processamento comum para dados categóricos.
- A codificação one-hot transforma os dados categóricos em uma representação vetorial binária.



Pré-processamento dos Dados

○ Normalização de dados

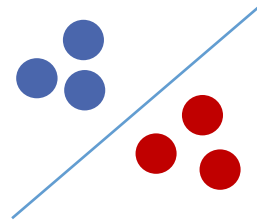
- Transformar os valores de diferentes variáveis para a mesma ordem de grandeza [0, 1] antes de serem utilizadas em algoritmos de ML;
- É uma prática para evitar que o algoritmo fique enviesado para as variáveis com maior ordem de grandeza;
- Normalização Min-Max:

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

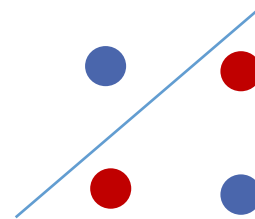
Algoritmos de Classificação

○ Redes Neurais Artificiais (RNA)

- São modelos computacionais inspirados no funcionamento do cérebro humano. Elas são compostas por unidades de processamento chamadas de neurônios artificiais, que estão organizados em camadas e interconectados por conexões ponderadas.
- Problemas não-linearmente separáveis.



(a) Linearmente separável



(b) Não-linearmente separável.

Algoritmos de Classificação

○ ***Multilayer Perceptron (MLP)***

- É um tipo de RNA formada por pelo menos três camadas: Uma camada de entrada, uma ou mais camadas ocultas, composta por neurônios interconectados e uma camada de saída.

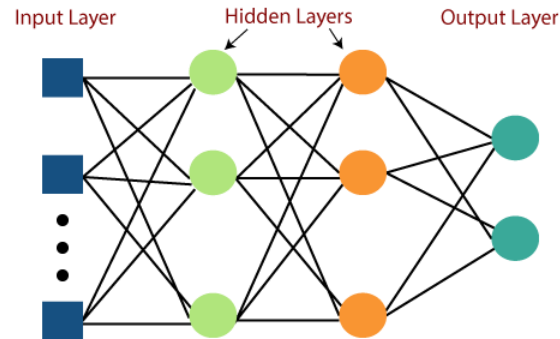


Figura - Arquitetura *Multilayer Perceptron* (MLP)

Algoritmos de Classificação

○ Neurônio Artificial

- Cada neurônio é responsável por realizar uma soma ponderada das saídas dos neurônios da camada anterior e aplicar uma função de ativação não-linear a essa soma ponderada.

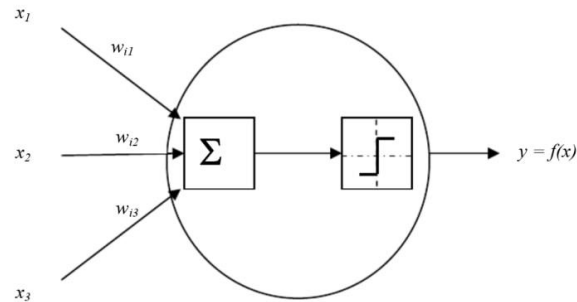
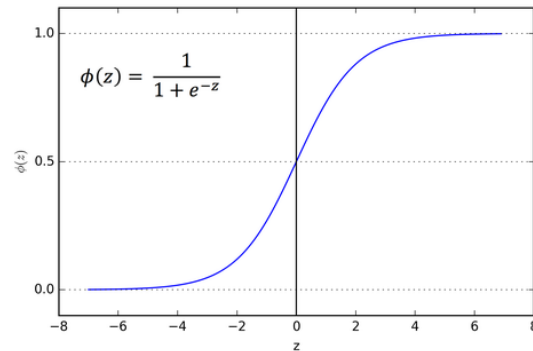


Figura – Representação do Neurônio Artificial

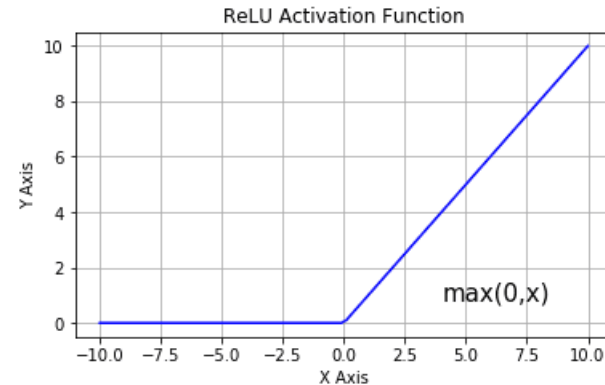
Algoritmos de Classificação

○ Função de Ativação

- Introduzem a não-linearidades nas saídas dos neurônios, permitindo que a MLP seja capaz de modelar relações complexas e não-lineares, aprenda padrões mais complexos e faça previsões mais precisas.



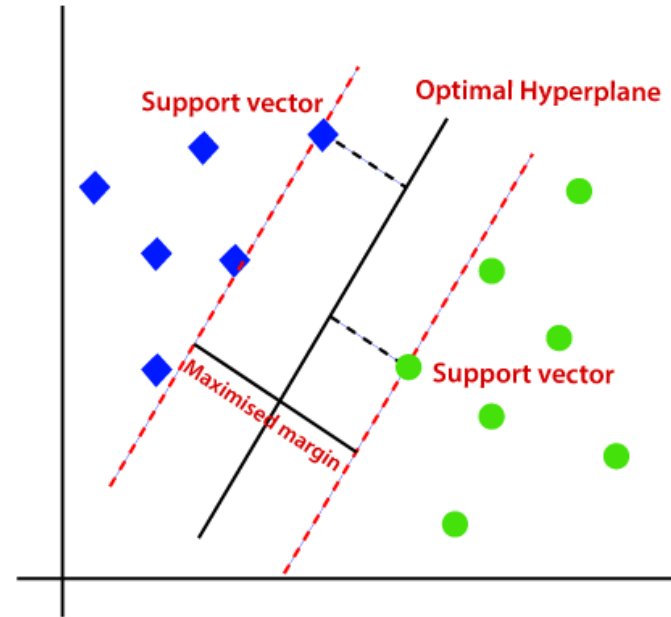
(A) Função de Ativação – Tangente Hiperbólica



(b) Função de Ativação – ReLU

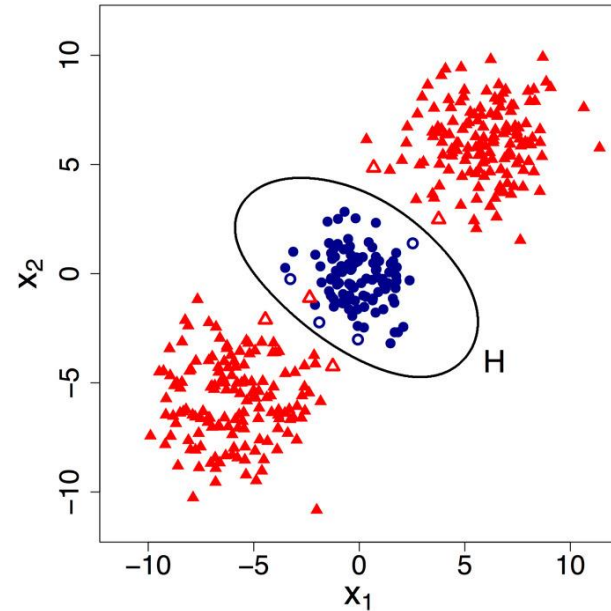
Algoritmos de Classificação

- **Support Vector Machine (SVM)**
 - Algoritmo de aprendizado de máquina supervisionado, cujo objetivo é encontrar um hiperplano que melhor separe as duas classes.
 - Para separar as duas classes, existem muitos hiperplanos possíveis. O objetivo é encontrar um plano que tenha a distância máxima entre os pontos de dados de ambas as classes.



Algoritmos de Classificação

- **Support Vector Machine (SVM)**
 - Em problemas não-linearmente separáveis usa-se o kernel. O kernel é uma função que mapeia os dados de entrada para um espaço de características de maior dimensionalidade, onde a separação linear pode ser mais facilmente alcançada.
 - Kernel linear, kernel polinomial, kernel gaussiano (RBF - Radial Basis Function).



Avaliação dos Classificadores

- **Matriz de Confusão:** Resume o desempenho de um modelo de classificação em relação aos dados de testes.

		Valor Predito	
		Sim	Não
Real	Sim	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Avaliação dos Classificadores

- **Acurácia (acc):** Número de previsões corretas de um modelo para o total de amostras:

$$\frac{VP + VN}{N}$$

- **Precisão:** Observações positivas classificadas corretamente entre todos aqueles preditos como positivos:

$$\frac{VP}{VP + FP}$$

Avaliação dos Classificadores

- **Recall:** Taxa de verdadeiros positivos:

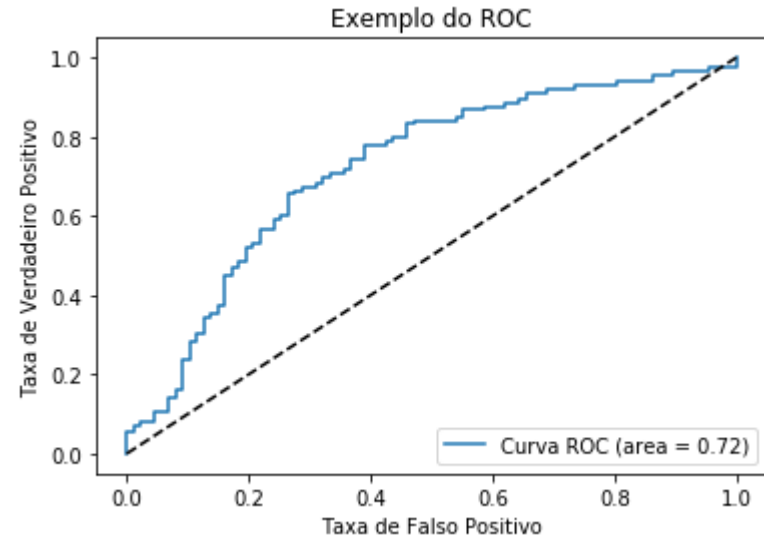
$$\frac{VP}{VP + FN}$$

- **F1-Score:** Define a média harmônica da precisão e recall:

$$\frac{2 \times precisao \times recall}{precisao + recall}$$

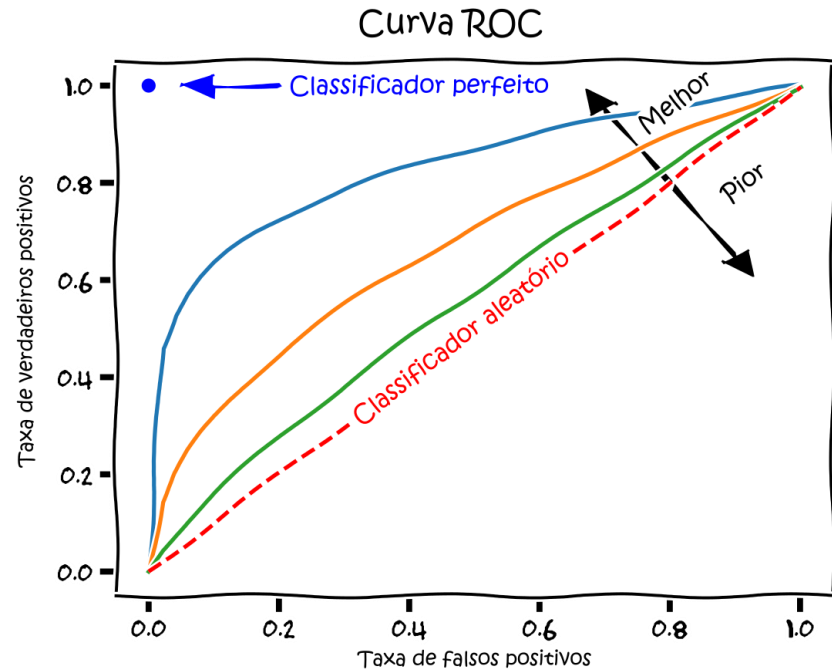
Avaliação dos Classificadores

- **Curva ROC:** Quanto mais próximo o ponto da curva ROC estiver do canto superior esquerdo ($TPR=1$ e $FPR=0$), melhor é a performance do classificador.



Avaliação dos Classificadores

- **AUC:** Representa a área sob a curva, variando de 0 a 1. Quanto maior a AUC, melhor é a capacidade do classificador em distinguir corretamente entre as classes.



Referências

- [1] Grus, Joel Data Science do Zero / Joel Grus; traduzido por Welington Nascimento. - Rio de Janeiro : Alta Books, 2016. 336 p. : il. ; 3,8 MB.
- [2] Scikit-Learn - <https://scikit-learn.org/stable/>

Obrigada!

- *Alguma dúvida?*
andressagomes@alu.ufc.br