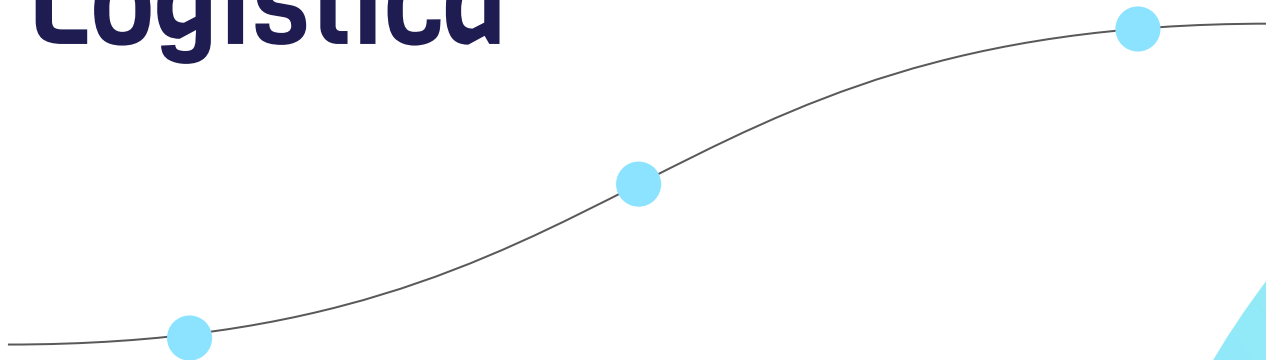


Regressão Logística



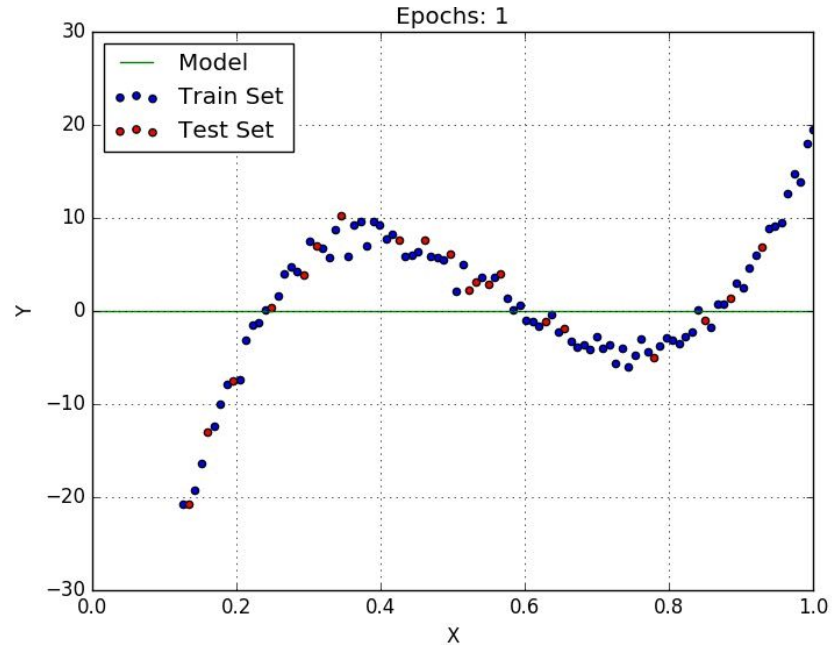
CCF 726

Matheus Freitas Martins (ES111281)

O que é? Como funciona?

A proposta principal do algoritmo é prever a probabilidade de ocorrência de um **evento binário**.

- Probabilidade entre **0** e **1**.



Probabilidade de
pertencer à classe
positiva.

$$p = \frac{1}{(1 + e^{(-z)})}$$

combinação linear
dos coeficientes e
das variáveis de
entrada.

$$z = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n$$

- **penalty**: 'l1' ou 'l2'
- **dual**: True ou False
- **tol**: Valor numérico (tolerância)
- **C**: Valor numérico (float)
- **fit_intercept**: True ou False
- **class_weight**: 'None' ou 'balanced'
- **random_state**: Valor numérico ou None
- **solver**: 'newton-cg', 'lbfgs', 'liblinear', 'sag' ou 'saga'
- **max_iter**: Valor numérico (inteiro)

Vantagens e Desvantagens

- É um algoritmo simples e eficiente que funciona bem na maioria dos casos, especialmente quando o conjunto de dados não é muito grande e as características são aproximadamente lineares.
- Funciona bem para problemas de classificação binária.
- É menos propenso a ser influenciado por dados com ruídos e pode evitar o overfitting.
- A saída pode ser interpretada como uma probabilidade.



- A regressão logística tende a ter um desempenho ruim quando existem limites de decisão não lineares.
- Não é adequado para grandes dimensões de dados.
- Pode não funcionar bem se os atributos são correlacionados.
- Pode não ser a melhor escolha se a relação entre atributos e classe é não-linear.



- Multinomial
- Linearidade
- Independência das observações
- Ausência de valores ausentes
- Tamanho da amostra
- Ausência de multicolinearidade perfeita

- Classificação Binária
- Problemas com variáveis categóricas
- Quando as variáveis independentes são lineares
- Problemas que requerem probabilidades

Exemplos: diagnóstico médico, risco de crédito, marketing, ciências sociais, spam de email, sentimento de análise de texto.

Ref.

[1] SCIKIT-LEARN. Logistic Regression. Disponível em:
https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html. Acesso em:
16 maio 2023.

[2] WIKIPÉDIA. Regressão logística. Disponível em:
https://pt.wikipedia.org/wiki/Regress%C3%A3o_log%C3%ADstica. Acesso em: 18 maio 2023.

[3] SPRINGBOARD. What Is Logistic Regression? Disponível em:
<https://www.springboard.com/blog/data-science/what-is-logistic-regression/>. Acesso em: 20 maio 2023.

OBRIGADO!

matheus.f.martins@ufv.br

