



**UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS JARDINS DE ANITA - ITAPAJÉ**

Dicionário: Features do Dataset Iris e Dataset Qualidade do Vinho

**Autores: Giovanna Dias Castro de Oliveira
Jônatas Fernandes Silva
Leandro Nascimento Adegas**

Disciplina: Aprendizado de Máquina e Reconhecimento de Padrões

Itapajé, Maio, 2025



Conteúdo

1	Iris-Dataset	3
1.1	O que é	3
1.2	Features	3
1.3	Significado de Cada Feature	3
1.3.1	Feature Id	3
1.3.2	Feature SepalLengthCm	4
1.3.3	Feature SepalWidthCm	4
1.3.4	Feature PetalLengthCm	4
1.3.5	Feature PetalWidthCm	4
1.3.6	Feature Species	4
2	Qualidade do Vinho Dataset	4
3	O que é	4
4	Features Vinho Branco e Vermelho	4
5	Referências	6



1 Iris-Dataset

1.1 O que é

O conjunto de dados Íris é um dos conjuntos de dados mais populares em ciência de dados e aprendizado de máquina. Ele contém informações sobre 150 flores de íris, divididas em três espécies diferentes: Iris setosa, Iris versicolor e Iris virginica. Cada flor tem quatro características medidas em centímetros: comprimento da sépala, largura da sépala, comprimento da pétala e largura da pétala.

Este conjunto de dados é frequentemente usado como um exemplo para demonstrar técnicas de análise de dados e classificação. Alguns dos usos comuns do conjunto de dados Iris incluem:

Classificação de flores: Com base nas características medidas das flores, podemos treinar um modelo de aprendizado de máquina para prever a espécie de uma flor desconhecida. Isso é um exemplo de classificação de dados.

Visualização de dados: Podemos usar as quatro características do conjunto de dados Íris para criar visualizações interessantes, como gráficos de dispersão e diagramas de caixa. Isso nos ajuda a entender melhor as relações entre as diferentes características e as espécies de íris.

Teste de algoritmos de aprendizado de máquina: Como o conjunto de dados Iris é relativamente pequeno e bem compreendido, ele é frequentemente usado para testar algoritmos de aprendizado de máquina. Isso nos ajuda a entender como diferentes algoritmos de aprendizado de máquina funcionam e como eles podem ser ajustados para obter melhores resultados.

Em resumo, o conjunto de dados Iris é uma ferramenta valiosa para a comunidade de ciência de dados e aprendizado de máquina. Ele nos permite explorar técnicas de análise de dados e classificação em um conjunto de dados bem compreendido e interessante [1].

1.2 Features

O dataset Iris é composto por seis (06) features, a saber:

1. Id;
2. SepalLengthCm;
3. SepalWidthCm;
4. PetalLengthCm;
5. PetalWidthCm; e
6. Species

1.3 Significado de Cada Feature

Cada feature agrupa uma série de valores no dataset. A seguir veremos os valores que são agrupados em cada feature.

1.3.1 Feature Id

A feature **Id** representa apenas a identificação de cada flor dentro do dataset, ou seja, ele é apenas um indicador para guiar o analista. Dessa forma, vem facilitar quanto a identificação de qual flor está sendo observada.



1.3.2 Feature SepalLengthCm

A feature **SepalLengthCm** representa o comprimento em centímetros das sépalas que compõem cada flor. A saber, sépalas são folhas modificadas, geralmente de coloração verde, localizadas abaixo das pétalas e com função de proteger o botão floral enquanto ainda não abriu [2].

1.3.3 Feature SepalWidthCm

A feature **SepalWidthCm** representa a largura em centímetros das sépalas que compõem cada flor.

1.3.4 Feature PetalLengthCm

A feature **PetalLengthCm** representa o comprimento em centímetros das pétalas que compõem cada flor. A saber, pétalas são uma unidade da corola, pode possuir cores vivas e marcantes e sua função é atrair agentes polinizadores. Inserem-se logo após as sépalas. [2].

1.3.5 Feature PetalWidthCm

A feature **PetalLengthCm** representa a largura em centímetros das pétalas que compõem cada flor.

1.3.6 Feature Species

A feature **Species** representa a espécie que cada flor indicada por seu **Id** pertence.

2 Qualidade do Vinho Dataset

3 O que é

Os dois conjuntos de dados estão relacionados com as variantes vermelha e branca do vinho português "Vinho Verde". Para mais detalhes, consulte a referência [3] [4].

4 Features Vinho Branco e Vermelho

O Dataset Qualidade do Vinho é composto por doze (12) features, a saber:

1. fixed acidity: nível de acidez fixa;
2. volatile acidity: nível de acidez volátil;
3. citric acid: nível de ácido cítrico;
4. residual sugar: nível de de açúcar residual;
5. chlorides: nível de cloreto;
6. free sulfur dioxide: nível de dióxido de enxofre livre;
7. total sulfur dioxide: nível total de dóxido de enxofre;
8. density: densidade (base h2o -j 1.0);



Prova 01

9. pH: O pH é uma escala numérica que determina o grau de acidez de uma solução aquosa, baseado na concentração de íons hidrônio (H_3O^+). Soluções ácidas possuem excesso de íons hidrônio e pH menor do que 7. Soluções básicas possuem excesso de íons hidroxila (OH^-) e valores de pH superiores a 7. Soluções consideradas neutras têm igual concentração de íons H_3O^+ e íons OH^- , e sua medida de pH é 7. [5];
10. sulphates: nível de sulfatos, a saber: Os sulfatos são compostos iônicos que contêm o ânion SO_4^{2-} , que é chamado de ânion sulfato. [6];
11. alcohol: nível de álcool presente no vinho; e
12. quality (score between 0 and 10): índice de qualidade do vinho



5 Referências

- [1] “O conjunto de dados Íris é um dos conjuntos de dados mais populares em ciência de dados e aprendizado de máquina..” <https://www.dio.me/articles/o-conjunto-de-dados-iris-e-um-dos-conjuntos-de-dados-mais-populares-em-ciencia> Online; Acesso em 2025-06-03.
- [2] “Partes da flor: verticilos florais, pedúnculo e pistilo.” <https://querobolsa.com.br/enem/biologia/flor-partes-da-planta>. Online; Acesso em 2025-06-03.
- [3] P. Cortez, A. Cerdeira, F. Almeida, T. Matos, and J. Reis, “Modeling wine preferences by data mining from physicochemical properties,” *Viticulture Commission of the Vinho Verde region (CVRVV)*, 4050-501 Porto, Portugal, 2009.
- [4] “Qualidade do vinho tinto.” <https://www.kaggle.com/datasets/uciml/red-wine-quality-cortez-et-al-2009>. Online; Acesso em 2025-06-03.
- [5] “O que é ph?.” <https://mundoeducacao.uol.com.br/quimica/voce-sabe-que-significa-ph-.htm>. Online; Acesso em 2025-06-03.
- [6] “Sulfatos.” <https://brasilescola.uol.com.br/quimica/sulfatos.htm>. Online; Acesso em 2025-06-03.