**Wine Data Set:**

Esses dados são resultados de uma análise química de vinhos cultivados na mesma região da Itália, mas derivados de três cultivares diferentes. A análise determinou as quantidades de 13 constituintes encontrados em cada um dos três tipos de vinhos. Acho que o conjunto de dados inicial tinha cerca de 30 variáveis, mas por algum motivo só tenho a versão 13 dimensional. Eu tinha uma lista do que eram as cerca de 30 variáveis, mas a.) Eu perdi, eb), não saberia quais 13 variáveis ​​estão incluídas no conjunto.

Os atributos são

1) Álcool

2) ácido málico

3) Cinza

4) Alcalinidade das cinzas

5) Magnésio

6) Fenóis totais

7) Flavanóides

8) Fenóis não flavanoides

9) Proantocianinas

10) Intensidade da cor

11) Matiz

12) OD280 / OD315 de vinhos diluídos

13) Proline

Nenhuma estatística disponível, mas sugere padronizar variáveis ​​para determinados usos (por exemplo, para nós com classificadores que NÃO são invariantes de escala) NOTA: o primeiro atributo é o identificador de classe (1-3)

**Hepatitis Data Set**

Os atributos são

Informação de Atributo:

1. Class: DIE, LIVE

2. AGE: 10, 20, 30, 40, 50, 60, 70, 80

3. SEX: male, female

4. STEROID: no, yes

5. ANTIVIRALS: no, yes

6. FATIGUE: no, yes

7. MALAISE: no, yes

8. ANOREXIA: no, yes

9. LIVER BIG: no, yes

10. LIVER FIRM: no, yes

11. SPLEEN PALPABLE: no, yes

12. SPIDERS: no, yes

13. ASCITES: no, yes

14. VARICES: no, yes

15. BILIRUBIN: 0.39, 0.80, 1.20, 2.00, 3.00, 4.00

-- see the note below

16. ALK PHOSPHATE: 33, 80, 120, 160, 200, 250

17. SGOT: 13, 100, 200, 300, 400, 500,

18. ALBUMIN: 2.1, 3.0, 3.8, 4.5, 5.0, 6.0

19. PROTIME: 10, 20, 30, 40, 50, 60, 70, 80, 90

20. HISTOLOGY: no yes

O atributo BILIRUBIN parece ser continuamente valorizado. Verifiquei isso com o doador, Bojan Cestnik, que respondeu: Sobre o banco de dados de hepatite e o problema de BILIRUBIN, gostaria de dizer o seguinte: BILIRUBIN é um atributo contínuo (= o número de seus "valores" no arquivo ASDOHEPA.DAT é negativo !!!); "valores" são citados porque, quando se fala sobre o atributo contínuo, não existem todos os valores possíveis. No entanto, eles representam os chamados valores "limites"; de acordo com esses valores "limites", o atributo pode ser discretizado. Ao mesmo tempo, por causa do atributo contínuo, pode-se realizar algum outro teste desde que a informação contínua seja preservada. Espero que essas linhas tenham respondido pelo menos aproximadamente à sua pergunta.

**Colab:**

**Wine Data Set:**

**01-ARVORE & knn**

[**https://colab.research.google.com/drive/1bnxjZLU1v0zO6X59xEwDuLy2\_boVTyp#scrollTo=UwrLXKJ7BZd2**](https://colab.research.google.com/drive/1bnxjZLU1v0zO6X59xEwDuLy2_boVTyp#scrollTo=UwrLXKJ7BZd2)

**01-KNN IMPROVE**

[**https://colab.research.google.com/drive/1iEbCrN4mEOFQ43ii9rT\_5-5-xPAMAH7F#scrollTo=YKVmVVEHUnjH**](https://colab.research.google.com/drive/1iEbCrN4mEOFQ43ii9rT_5-5-xPAMAH7F#scrollTo=YKVmVVEHUnjH)

**Hepatitis Data Set**

**02-ARVORE & knn**

[**https://colab.research.google.com/drive/1LagyGOtQGBm6D4TRadkeADhpTxgO2C7k#scrollTo=xyBhA4qLsDj7**](https://colab.research.google.com/drive/1LagyGOtQGBm6D4TRadkeADhpTxgO2C7k#scrollTo=xyBhA4qLsDj7)

**02- KNN IMPROVE**

[**https://colab.research.google.com/drive/1wZ2gNizDuoA0O\_KCpwXMwgzgG2PBlIsu#scrollTo=5jGUzYwJeaQp**](https://colab.research.google.com/drive/1wZ2gNizDuoA0O_KCpwXMwgzgG2PBlIsu#scrollTo=5jGUzYwJeaQp)

|  |  |  |
| --- | --- | --- |
|  | **Wine Data Set** | **Hepatitis Data Set** |
| **Árvore-entropy** | **89%** | **88%** |
| **Árvore-gini** | **89%** | **94%** |
| **Knn-euclidean N10** | **50%** | **81%** |
| **Knn-euclidean N20** | **58%** | **81%** |
| **Knn-euclidean N30** | **58%** | **81%** |
| **Knn-chebyshev N10** | **50%** | **81%** |
| **Knn-chebyshev N20** | **58%** | **81%** |
| **Knn-chebyshev N30** | **56%** | **81%** |
| **Knn-improve K5** | **56%** | **81%** |
| **Knn-improve k10** | **56%** | **81%** |

**GitHub:**

[**https://github.com/jailcomfranssa/Projeto--Arvores-de-decisao-kNN**](https://github.com/jailcomfranssa/Projeto--Arvores-de-decisao-kNN)**.**

**Aluno:** Jailson Silva de França

**Matricula:**20182370048

**Aluno:** Helder Jerônimo Leite Rangel

**Matricula:** 20161370009