

4º Lista de Inteligência Artificial - Iasmin Ferreira e Oliveira

LINKS: ID3: <https://github.com/iasminfeo/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/ID3.PY>

Cart: <https://github.com/iasminfeo/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/CART.py>

C45: <https://github.com/iasminfeo/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/C45.py>

Seção 01

Código:

```
from __future__ import annotations
from typing import Tuple
import numpy as np
import pandas as pd

def train_test_split_stratified(y: np.ndarray, test_size: float = 0.2, seed: int = 42) -> Tuple:
    rng = np.random.default_rng(seed)
    idx = np.arange(len(y))
    test_idx = []
    for c in np.unique(y):
        class_idx = idx[y == c]
        rng.shuffle(class_idx)
        n_test = max(1, int(round(test_size * len(class_idx))))
        test_idx.extend(class_idx[-n_test:])
    test_idx = np.array(sorted(test_idx))
    train_idx = np.array([i for i in idx if i not in test_idx])
    return train_idx, test_idx

def discretize_equal_frequency(series: pd.Series, bins: int = 4, labels: bool = True) -> pd.Series:
    """Discretiza por quantis (-mesmo número de amostras por faixa)."""
    q = np.linspace(0, 1, bins + 1)
    edges = np.unique(series.quantile(q).values)
    edges[0] = -np.inf
    edges[-1] = np.inf
    cats = pd.cut(series, bins=edges, include_lowest=True)
    return cats.astype(str) if labels else cats

def discretize_equal_width(series: pd.Series, bins: int = 4, labels: bool = True) -> pd.Series:
    """Discretiza por largura fixa (intervalos iguais)."""
    cats = pd.cut(series, bins=bins, include_lowest=True)
    return cats.astype(str) if labels else cats

def impute_single(df: pd.DataFrame) -> pd.DataFrame:
    """Impute NaNs: numéricos -> mediana; categóricas/objetos -> moda."""
    out = df.copy()
    for col in out.columns:
        if out[col].dtype == object:
            if out[col].isna().any():
                out[col] = out[col].fillna(out[col].mode().iloc[0])
        else:
            if out[col].isna().any():
                out[col] = out[col].fillna(out[col].median())
    return out
```

Seção 2

ID3 (ganho de informação; atributos categóricos)

<https://github.com/iasminfeio/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/ID3.PY>

C45 (razão de ganho; contínuos por limiar; categórico multi-ramo)

<https://github.com/iasminfeio/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/CART.py>

CART (índice de Gini; divisões binárias; categórico por subconjunto)

<https://github.com/iasminfeio/IA/blob/main/Listas-c%C3%B3digos/Lista4/Code/C45.py>

Seção 3

Configuração: *split* estratificado 80/20 (`seed=42`), max_depth=6. 3.3) ID3 Acurácia (treino): 0.8808 \nAcurácia (teste): 0.8034 Matriz de confusão (teste): labels: 0, 1 ||0|1|
|-|-| |0|99|11| |1|24|44| Árvore:

```
[Sex]
-> female:
[Folha]
-> 1:
[Face]
-> (14.454, 30.5):
[Embarked]
-> C:
[Age]
-> (35.0, inf]:
[Parch]
-> 0:
Folha: 0
-> S:
Folha: 1
-> (30.5, inf]:
Folha: 1
-> 2:
[Age]
-> (-inf, 22.0]:
Folha: 1
-> (22.0, 28.0]:
[Parch]
-> 0:
[Embarked]
-> C:
Folha: 1
-> S:
[Face]
-> (14.454, 30.5):
Folha: 1
-> (30.5, inf]:
Folha: 1
-> (7.896, 14.454]:
Folha: 1
-> 1:
[SibSp]
-> 0:
Folha: 1
-> 1:
[Embarked]
-> S:
Folha: 0
-> 2:
Folha: 1
-> 2:
Folha: 1
-> 3:
Folha: 1
-> (28.0, 35.0):
Folha: 1
-> (35.0, inf]:
[Parch]
-> 0:
[Face]
-> (14.454, 30.5):
[Embarked]
-> S:
```

3.2) C45 Acurácia (treino): 0.8219 Acurácia (teste): 0.8202 Matriz de confusão (teste):
labels: 0, 1 ||0|1| |-|-| |0|103|7| |1|25|43| Árvore:

```

[Sex]
=> female:
  [trip < 4]
    => < 1
      [Pclass < 2.5]
        => < 1
          [Fare < 28.8562]
            => < 1
              [Age < 52.5]
                => < 1
                  [Paroh < 1.5]
                    => < 1
                      Folha: 1
                    => >= 1
                      Folha: 1
                    => >= 1
                      [Embarked]
                        => S:
                          Folha: 1
                        => >= 1
                          Folha: 1
                        => >= 1
                          [Fare < 32.8813]
                            => < 1
                              [Age < 1.5]
                                => < 1
                                  Folha: 1
                                => >= 1
                                  [Embarked]
                                    => C:
                                      Folha: 1
                                    => Q:
                                      Folha: 1
                                    => S:
                                      Folha: 0
                                  => >= 1
                                    Folha: 0
                                  => >= 1
                                    Folha: 0
                                => >= 1
                                  Folha: 0
                              => >= 1
                                Folha: 0
                            => >= 1
                              Folha: 1
                              => >= 1
                                [Fare < 387.665]
                                  => < 1
                                    [trip < 4.5]
                                      => < 1
                                        [Paroh < 2.5]
                                          => < 1
                                            [Pclass < 1.5]
                                              => < 1
                                                Folha: 0
                                              => >= 1
                                                Folha: 0
                                              => >= 1
                                                Folha: 0
                                            => >= 1
                                              Folha: 0
                                          => >= 1
                                            Folha: 0
                                        => >= 1
                                          Folha: 0
                                      => >= 1
                                        Folha: 0
                                    => >= 1
                                      Folha: 0
                                  => >= 1
                                    Folha: 0
                                => >= 1
                                  Folha: 0
                              => >= 1
                                Folha: 0
                            => >= 1
                              Folha: 0
                          => >= 1
                            Folha: 0
                        => >= 1
                          Folha: 0
                      => >= 1
                        Folha: 0
                    => >= 1
                      Folha: 0
                  => >= 1
                    Folha: 0
                => >= 1
                  Folha: 0
              => >= 1
                Folha: 0
            => >= 1
              Folha: 0
          => >= 1
            Folha: 0
        => >= 1
          Folha: 0
      => >= 1
        Folha: 0
    => >= 1
      Folha: 0
  => >= 1
    Folha: 0
=> male:
  [Age < 1.5]
    => < 1
      Folha: 1
    => >= 1
      [Fare < 387.665]
        => < 1
          [trip < 4.5]
            => < 1
              [Paroh < 2.5]
                => < 1
                  [Pclass < 1.5]
                    => < 1
                      Folha: 0
                    => >= 1
                      Folha: 0
                    => >= 1
                      Folha: 0
                  => >= 1
                    Folha: 0
                => >= 1
                  Folha: 0
              => >= 1
                Folha: 0
            => >= 1
              Folha: 0
          => >= 1
            Folha: 0
        => >= 1
          Folha: 0
      => >= 1
        Folha: 0
    => >= 1
      Folha: 0
  => >= 1
    Folha: 0

```

3.3) CART Acurácia (treino): 0.8682 Acurácia (teste): 0.8258 Matriz de confusão (teste):
 labels: 0, 1 |0|1| |-|-|-| |0|100|10| |1|21|47| Árvore

```

[Sex ~ ('female')]
=> ~ 1
[Polaco < 2.5]
=> < 1
[Fare < 28.8562]
=> < 1
[Fare < 28.2312]
=> < 1
[Age < 53.5]
=> < 1
[Stbdp < 0.5]
=> < 1
Folha: 1
=> >=;
Folha: 1
=> >=;
[Polaco < 1.5]
=> < 1
Folha: 1
=> >=;
Folha: 0
=> >=;
Folha: 0
=> >=;
Folha: 1
=> >=;
[Fare < 20.6625]
=> < 1
[Age < 7]
=> < 1
Folha: 1
=> >=;
[Fare < 8.0396]
=> < 1
[Age < 29.25]
=> < 1
Folha: 1
=> >=;
Folha: 0
=> >=;
[Fare < 15.8]
=> < 1
Folha: 0
=> >=;
Folha: 1
=> >=;
[Parch < 0.5]
=> < 1
Folha: 1
=> >=;
[Age < 5.5]
=> < 1
[Age < 3.5]
=> < 1
Folha: 0
=> >=;
Folha: 1

```