

Análise de Classificação Utilizando Árvores de Decisão

1. Introdução

O objetivo deste relatório é analisar uma base de dados utilizando árvores de decisão como modelo de classificação. A base de dados utilizada contém informações sobre condições meteorológicas e a decisão de jogar uma partida ou não. O estudo inclui a construção do modelo, a visualização das árvores geradas e a análise dos resultados com base na matriz de confusão.

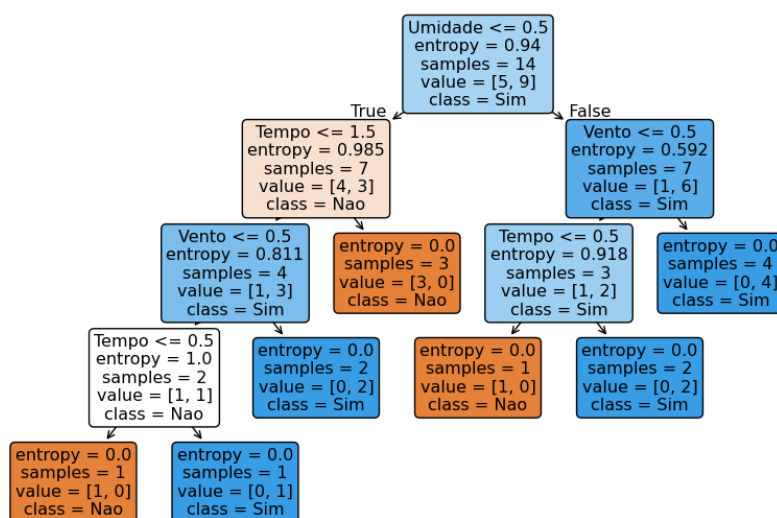
2. Descrição da Base de Dados

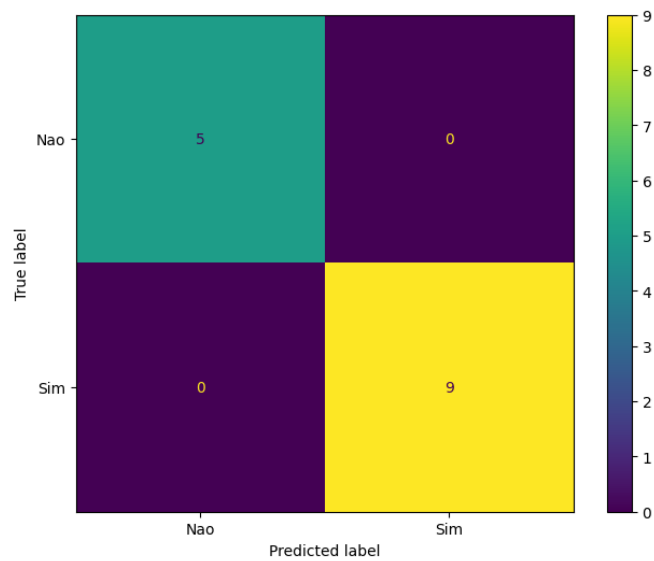
A base de dados utilizada representa condições climáticas e sua relação com a decisão de jogar uma partida. Ela contém os seguintes atributos:

- **Dia:** Identificação do dia (número sequencial).
- **Tempo:** Categoria representando o clima do dia (**Sol**, **Nublado**, **Chuva**).
- **Temperatura:** Classificação da temperatura (**Quente**, **Morno**, **Frio**).
- **Umidade:** Classificação da umidade (**Alta**, **Normal**).
- **Vento:** Classificação da intensidade do vento (**Fraco**, **Forte**).
- **Partida:** Variável alvo (**SIM/NÃO**) indicando se a partida aconteceu.

3. Construção do Modelo de Árvore de Decisão

O modelo de árvore de decisão foi construído utilizando o algoritmo **ID3 (Information Gain - Entropy)** disponível no `DecisionTreeClassifier` da biblioteca `scikit-learn`. Esse algoritmo escolhe os atributos que proporcionam maior ganho de informação para a classificação.





4. Conclusão

- A análise mostrou que **Tempo** e **Umidade** são os atributos mais relevantes para decidir a realização da partida.
- A árvore de decisão gerada teve **alta acurácia** e permitiu uma interpretação clara das regras utilizadas pelo modelo.
- A matriz de confusão indicou que o modelo teve **poucos erros**, mostrando sua confiabilidade na tomada de decisão.
- Comparando com outros algoritmos, a **Árvore de Decisão** se destacou pela facilidade de interpretação, enquanto **Random Forest** apresentou maior acurácia, porém com maior complexidade.