

Proposta - Trabalho Prático 3

Classificação das diversas classes de animais de um zoológico

Base de dados e contextualização

A base de dados escolhida foi encontrada no Kaggle, denominada como Car Features and MSRP (<https://www.kaggle.com/uciml/zoo-animal-classification>). A base de dados está no formato CSV e possui 101 animais diferentes descritos por 16 atributos. O dataset foi fornecido pelo website UCI Machine Learning, pelo criador Richard Forsyth. Dentre os atributos, temos o nome do animal, se o animal tem cabelo, se o animal tem penas, se o animal é aquático, se o animal é predador, etc. Os animais são classificados em 7 classes distintas:

1. Mamífero
2. Passáro
3. Réptil
4. Peixe
5. Anfíbio
6. Inseto
7. Invertebrado

A base de dados foi escolhida por possuir uma documentação clara e atributos de fácil entendimento para o treinamento de dados nos algoritmos de classificação, como o Multilayer Perceptron, por exemplo.

Objetivo

Treinar um modelo de classificação probabilístico através da base de dados citada na seção anterior. A princípio, o objetivo principal é utilizar o algoritmo de classificação Multilayer Perceptron para o treinamento e validação do modelo, observando as acurácias obtidas para cada uma das 7 classes e tentando melhorar através de ajustes nos parâmetros do algoritmo. A base de dados também será treinada e avaliada no algoritmo de árvore de decisão comparando e discutindo os resultados obtidos com o algoritmo Multilayer Perceptron.