**Dados:**

A fonte de dados consistia em uma coleção de informações qualitativas e não veridicas sobre automóveis.

*preco:* preço do carro (muito alto, alto, medio, baixo)

*manutencao:* preço de manutenção (muito alto, alto, medio, baixo)

*portas:* quantidade de portas (2, 3, 4, 5mais)

*pessoas:* quantidade de portas (2, 4, 5mais)

*bagageiro:* tamanho do bagageiro (grande, medio, pequeno)

*seguranca:* segurança do carro (alta, media, baixa)

*aceitabilidade (atributo-alvo):* nível de aceitabilidade (inaceitavel, aceitavel, bom, muito bom)

**Análise dos Dados:**

Foi criado alguns modelos de classificação usando Árvores de Decisão (*DecisionTreeClassifier*) da biblioteca scikit-learn no Python.

Usando a função *train\_test\_split* separou-se 10% dos dados de maneira aleatória para treinamento do modelo de clasificação, e os demais dados para teste de acurácia.

Para entender como funciona e quais impactos traria ao resultado final o parâmetro *max\_depth* (ou em tradução livre "profundidade máxima") da função *DecisionTreeClassifier*. Foi alterado os valores, variando entre 4 e 12, sendo este o valor observado quando deixamos o sistema livre.

Para entender o impacto da variação deste parâmetro, foi observado a acurácia da classificação através da da função *accuracy\_score*. Dados apresentados na tabela e gráfico:

|  |  |
| --- | --- |
| **max\_depth** | **Acurácia (%)** |
| 4 | 87,28 |
| 5 | 90,17 |
| 6 | 92,49 |
| 7 | 93,06 |
| 8 | 96,53 |
| 9 | 96,53 |
| 10 | 99,42 |
| 11 | 97,69 |
| 12 | 97,69 |



**Comentários:**

O modelo de classificação por árvore de decisão apresentou bons resultados, com acurácia maior que 97% em cenário de *max\_depth = None*.

Terminar.....