

Aprendizado de Máquina Supervisionado

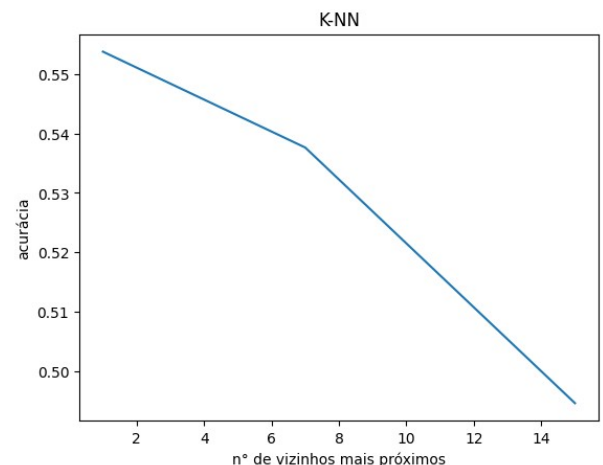
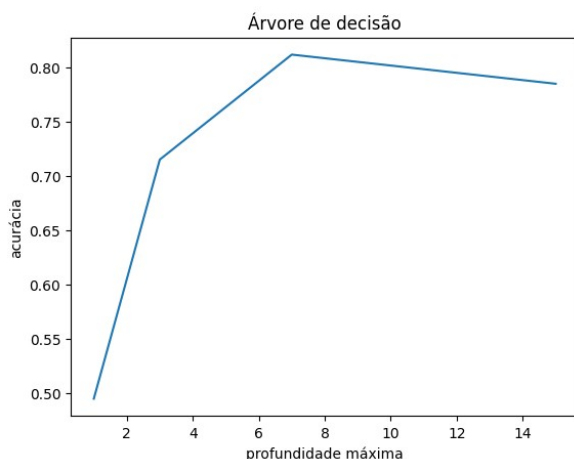
Relatório da atividade sobre Árvore de Decisão

Vinicius Cordeiro Fonseca
20250082351
UFRN/IMD
23/09/2025

Sobre os dados utilizados: Baixamos o banco de dados 4scenes e utilizamos os descritores de Haralick para extrair as características das imagens e armazená-las em vetores.

1) foi dividido o banco de dados em treinamento(70%), validação(15%) e teste (15%)

2) Utilizamos o conjunto de validação para selecionar os melhores hiperparâmetros para cada modelo, K sendo o número de vizinhos mais próximos e max_depth sendo a profundidade da árvore, e obtivemos os seguintes resultados:



3) em seguida treinamos os modelos com a união dos subconjuntos Treinamento + Validação utilizando os hiperparâmetros $K = 2$ e $\text{max_depth} = 7$.

4) calculamos as performances do K-NN e da Árvore de decisão utilizando a função `classification_report` e observamos também suas matrizes de confusão, como mostramos respectivamente a seguir. Além da acurácia, precisão e revocação serem melhores na Árvore de decisão, observamos também na horizontal principal da matriz de confusão que

houve mais acertos. Concluimos assim que, nesse caso, a Árvore de decisão é o melhor modelo entre eles.

```
[INFO] treinando modelo...
[INFO] avaliando...
      precision    recall  f1-score   support

   coast      0.42      0.63      0.51        49
   forest      0.67      0.70      0.68        56
 highway      0.50      0.34      0.41        41
   street      0.59      0.40      0.48        40

 accuracy          0.54      186
 macro avg      0.55      0.52      0.52      186
 weighted avg    0.55      0.54      0.53      186

[[31  4 10  4]
 [ 9 39  2  6]
[22  4 14  1]
[11 11  2 16]]
```

```
[INFO] treinando modelo...
[INFO] avaliando...
      precision    recall  f1-score   support

   coast      0.69      0.78      0.73        49
   forest      0.87      0.86      0.86        56
 highway      0.82      0.68      0.75        41
   street      0.69      0.72      0.71        40

 accuracy          0.77      186
 macro avg      0.77      0.76      0.76      186
 weighted avg    0.77      0.77      0.77      186

[[38  3  4  4]
 [ 3 48  0  5]
 [ 8  1 28  4]
 [ 6  3  2 29]]
```