

UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
INSTITUTO METROPOLE DIGITAL
IMD1101 - Aprendizado de Máquina – 2025.2
Técnica Supervisionada (DT)

Visando exercitar os conceitos aprendidos nas aulas de Aprendizado de Máquina Supervisionado, execute experimentos com o algoritmo Árvore de Decisão (DecisionTreeClassifier - scikit-learn), de tal forma que você possa escolher diferentes valores para dois parâmetros (*criterion* e *max_depth*) pertencentes à técnica estudada.

Os demais parâmetros serão utilizados com valores default. Dessa forma, você deverá criar um pequeno *GridSearch* com valores representados na Figura 1.

```
## Montando um GridSearch
par_criterion = ['gini', 'entropy', 'log_loss']
par_max_depth = [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]

for i in par_criterion:
    for j in par_max_depth:
```

Figura 1. Valores a serem utilizados para os parâmetros *criterion* e *max_depth*.

Como serão gerados 42 resultados de F1 score (3 valores de *criterion* x 14 valores de *max_depth*), selecione os dez (10) melhores valores (F1 Score) e os valores dos parâmetros correspondentes (*criterion* e *max_depth*). Para mais detalhes veja a Figura 2.

F1_score: 0.750 - gini - max_depth = 2	F1_score: 0.768 - entropy - max_depth = 2	F1_score: 0.768 - log_loss - max_depth = 2
F1_score: 0.750 - gini - max_depth = 3	F1_score: 0.768 - entropy - max_depth = 3	F1_score: 0.768 - log_loss - max_depth = 3
F1_score: 0.768 - gini - max_depth = 4	F1_score: 0.782 - entropy - max_depth = 4	F1_score: 0.782 - log_loss - max_depth = 4
F1_score: 0.764 - gini - max_depth = 5	F1_score: 0.773 - entropy - max_depth = 5	F1_score: 0.778 - log_loss - max_depth = 5
F1_score: 0.711 - gini - max_depth = 6	F1_score: 0.754 - entropy - max_depth = 6	F1_score: 0.754 - log_loss - max_depth = 6
F1_score: 0.703 - gini - max_depth = 7	F1_score: 0.785 - entropy - max_depth = 7	F1_score: 0.785 - log_loss - max_depth = 7
F1_score: 0.717 - gini - max_depth = 8	F1_score: 0.776 - entropy - max_depth = 8	F1_score: 0.776 - log_loss - max_depth = 8
F1_score: 0.680 - gini - max_depth = 9	F1_score: 0.716 - entropy - max_depth = 9	F1_score: 0.732 - log_loss - max_depth = 9
F1_score: 0.701 - gini - max_depth = 10	F1_score: 0.736 - entropy - max_depth = 10	F1_score: 0.736 - log_loss - max_depth = 10
F1_score: 0.686 - gini - max_depth = 11	F1_score: 0.752 - entropy - max_depth = 11	F1_score: 0.749 - log_loss - max_depth = 11
F1_score: 0.698 - gini - max_depth = 12	F1_score: 0.707 - entropy - max_depth = 12	F1_score: 0.752 - log_loss - max_depth = 12
F1_score: 0.671 - gini - max_depth = 13	F1_score: 0.732 - entropy - max_depth = 13	F1_score: 0.739 - log_loss - max_depth = 13
F1_score: 0.693 - gini - max_depth = 14	F1_score: 0.719 - entropy - max_depth = 14	F1_score: 0.736 - log_loss - max_depth = 14
F1_score: 0.695 - gini - max_depth = 15	F1_score: 0.694 - entropy - max_depth = 15	F1_score: 0.714 - log_loss - max_depth = 15

Figura 2. Resultados de F1 score para os parâmetros *criterion* e *max_depth*.

Depois de selecionados os dez (10) melhores valores de F1 score utilizando o Holdout (70/30), execute os mesmos experimentos agora utilizando o 10 fold cv. Na sequência, coloque os valores dos resultados na planilha, conforme a Figura 3.

DecisionTreeClassifier												
#	Base	Treinamento/Teste	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Conf. 5	Conf. 6	Conf. 7	Conf. 8	Conf. 9	Conf. 10
	Metodologias		F1-scr	F1-scr	F1-scr	F1-scr	F1-scr	F1-scr	F1-scr	F1-scr	F1-scr	F1-scr
1	HOG_128_20x20 (900 atributos)	10-fold CV 70/30										
2	HOG_128_24x24 (576 atributos)	10-fold CV 70/30										
3	HOG_128_18x18 (1.296 atributos)	10-fold CV 70/30										
4	HOG_128_30x30 (324 atributos)	10-fold CV 70/30										
5	HOG_64_16x16 (324 atributos)	10-fold CV 70/30										
6	HOG_64_32x32 (36 atributos)	10-fold CV 70/30										
7	HOG_128_30x30 PCA-75% (34)	10-fold CV 70/30										
8	HOG_128_32x32 (324 atributos)	10-fold CV 70/30										
9	HOG_256_32x32 (1764 atributos)	10-fold CV 70/30										
10	HOG_128_16x16 (1.764 atributos)	10-fold CV 70/30										
11	HOG_128_16x16 PCA-75% (115)	10-fold CV 70/30										
12	LBP_256_6R (50 atributos)	10-fold CV 70/30										
Média =>			#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
Desv. Pad. =>			#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
Conf. = criterion + max_depth												

Figura 3. Planilha com valores de F1 score para todas as bases.

Bom trabalho!