

# Classificação no Contexto de Telecomunicações

---

18 DE ABRIL

---

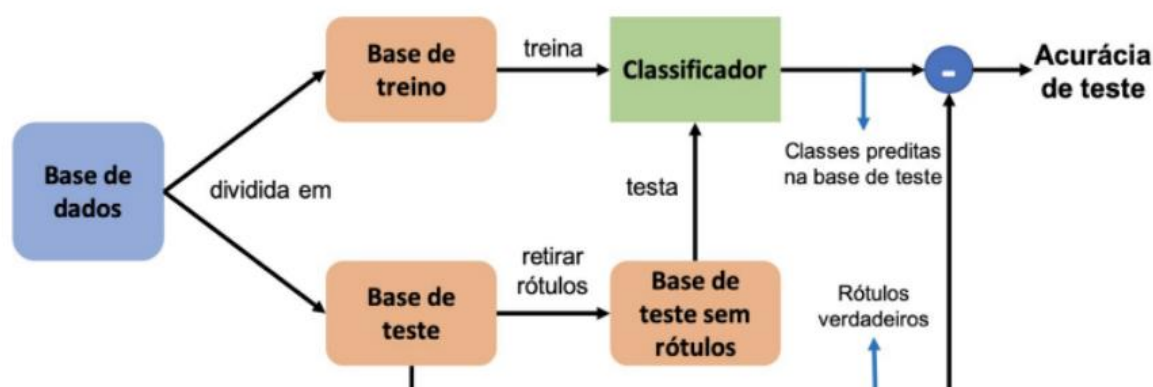
Criado por: Rafael Castro, Caio Damião, Álydson  
Lustoza, Luiz Jr



Nome do  
logotipo

## MÉTODO PARA ESTIMAR O DESEMPENHO

- ✚ Usando dois subconjuntos disjuntos: a base de treino e a base de teste.
- ✚ A base de treino é submetida ao modelo (classificador) para que seus parâmetros sejam calibrados de acordo com os dados apresentados. Após esta etapa, ocorre a etapa de predição de classes: os exemplos da base de teste são apresentados para o modelo treinado para que este realize a predição de suas classes. Medimos a qualidade do modelo comparando as classes preditas com as classes verdadeiras da base de teste, permitindo medir sua habilidade em classificar corretamente exemplos não vistos durante o treinamento. Este fluxo pode ser resumido pela figura a seguir.



*Fluxo resumido de problemas de Classificação. Fonte: (ESCOVEDO & KOSHIYAMA, 2020)*

- ✚ Existem diversas medidas para estimar o desempenho de um classificador, sendo a acurácia uma das mais utilizadas, que representa o percentual de acertos do classificador. Outra métrica bastante utilizada para problemas de classificação é a matriz de confusão, que oferece um detalhamento do desempenho do modelo de classificação, mostrando, para cada classe, o número de classificações corretas em relação ao número de classificações preditas pelo modelo.

## RESULTADOS

- ✚ **Teste com canal Crazy:** Valores calculados de acurácias e SER para diferentes exemplos de treino e teste.
- ✚ O melhor classificador é rede neural MultiLayer Perceptron (MLP) .

ACURÁCIA - CRAZY												
Treino/Teste		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Cálculo	Desvio
		600/5400	1200/4800	1800/4200	2400/3600	3000/3000	3600/2400	4200/1800	4800/1200	5400/600	Média	Padrão
Decision Tree	Treino	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	0
	Teste	97,04%	97,56%	97,88%	98,03%	98,07%	97,58%	97,83%	97,92%	97,33%	97,69%	0,00344529
Random Forest	Treino	100,00%	100,00%	100,00%	100,00%	99,93%	99,94%	100,00%	99,98%	99,91%	99,97%	0,00036401
	Teste	97,48%	97,92%	98,57%	98,14%	98,73%	98,12%	98,61%	98,67%	98,00%	98,25%	0,00423097
KNN	Treino	99,17%	99,25%	99,11%	99,00%	99,00%	99,08%	99,19%	99,25%	98,89%	99,10%	0,00123704
	Teste	98,22%	98,25%	98,79%	98,67%	98,90%	98,67%	98,61%	98,83%	98,50%	98,60%	0,00241253
Linear SVM	Treino	65,50%	68,33%	78,39%	70,17%	74,93%	74,75%	70,86%	75,31%	73,83%	72,45%	0,04028203
	Teste	61,37%	67,15%	77,24%	68,22%	73,80%	74,33%	71,28%	74,25%	71,50%	71,02%	0,04796392
SVM Kernel	Treino	98,67%	98,58%	98,89%	98,88%	98,73%	98,78%	98,83%	98,94%	98,59%	98,77%	0,00131825
	Teste	98,63%	98,33%	98,95%	98,69%	99,03%	98,67%	98,78%	98,75%	99,00%	98,76%	0,00218486
MLP	Treino	99,33%	98,83%	99,22%	98,92%	98,93%	98,81%	98,95%	98,88%	98,80%	98,96%	0,00186548
	Teste	98,72%	98,52%	98,83%	98,69%	99,10%	98,71%	98,83%	98,75%	99,00%	98,79%	0,00172852
Naive Bayes	Treino	99,50%	98,92%	99,06%	99,00%	98,87%	98,75%	98,98%	98,90%	98,70%	98,96%	0,00231415
	Teste	98,41%	98,60%	98,98%	98,72%	99,13%	98,75%	98,83%	98,83%	98,83%	98,79%	0,00207183

**Tabela de acurácia**

SER - CRAZY												
Decision Tree	SER	2,96%	2,44%	2,12%	1,97%	1,93%	2,42%	2,17%	2,08%	2,67%	2,31%	0,00344529
Random Forest	SER	2,52%	2,08%	1,43%	1,86%	1,27%	1,88%	1,39%	1,33%	2,00%	1,75%	0,00423097
Knn	SER	1,78%	1,75%	1,21%	1,33%	1,10%	1,33%	1,39%	1,17%	1,50%	1,40%	0,00241253
SVM Linear	SER	38,63%	32,85%	22,76%	31,78%	26,20%	25,67%	28,72%	25,75%	28,50%	28,98%	0,04796392
SVM kernel	SER	1,37%	1,67%	1,05%	1,31%	0,97%	1,33%	1,22%	1,25%	1,00%	1,24%	0,00218486
MLP	SER	1,28%	1,48%	1,17%	1,31%	0,90%	1,29%	1,17%	1,25%	1,00%	1,21%	0,00172852
Naive Bayes	SER	1,59%	1,40%	1,02%	1,28%	0,87%	1,25%	1,17%	1,17%	1,17%	1,21%	0,00207183

**Tabela da symbol error rate - SER**

- ✚ **Observação<sub>1</sub>:** Foram levados em consideração a média e o desvio padrão de todos os resultados.
- ✚ **Observação<sub>2</sub>:** Célula em amarelo é o classificador escolhido e em alaranjado é o aproximado.

✚ **Teste com o canal AWGN:** Foram calculados os valores de acurácias e SER para diferentes exemplos de treino e teste.

✚ Segundo os dados obtidos, o melhor classificador é Naive Bayes

ACURÁCIA - AWGN												
Treino/Teste		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Cálculo	Desvio
		600/5400	1200/4800	1800/4200	2400/3600	3000/3000	3600/2400	4200/1800	4800/1200	5400/600	Média	Padrão
Decision Tree	Treino	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	0
	Teste	97,15%	97,42%	97,71%	97,72%	96,77%	97,04%	97,89%	97,33%	96,50%	97,28%	0,004631
Random Forest	Treino	100,00%	100,00%	100,00%	99,96%	100,00%	99,92%	99,98%	99,96%	99,98%	99,98%	0,0002728
	Teste	97,74%	97,75%	98,07%	97,94%	97,47%	97,58%	97,78%	97,58%	97,50%	97,71%	0,0020167
KNN	Treino	97,67%	98,42%	98,72%	98,58%	98,73%	98,50%	98,83%	98,52%	98,46%	98,49%	0,0033789
	Teste	97,93%	97,75%	97,88%	97,94%	97,37%	97,71%	97,94%	97,42%	97,83%	97,75%	0,0021873
Linear SVM	Treino	88,00%	85,25%	94,39%	93,83%	93,37%	94,00%	94,57%	95,31%	94,31%	92,56%	0,0347466
	Teste	88,17%	82,73%	94,36%	94,47%	92,03%	94,29%	94,33%	94,33%	94,83%	92,17%	0,0412493
SVM Kernel	Treino	97,83%	98,25%	98,28%	98,08%	98,30%	98,00%	98,57%	98,35%	97,72%	98,15%	0,0026935
	Teste	98,39%	97,94%	98,36%	98,36%	97,57%	98,08%	98,50%	97,75%	97,67%	98,07%	0,0035087
MLP	Treino	97,67%	98,42%	98,56%	98,21%	98,40%	98,28%	98,45%	98,38%	97,81%	98,24%	0,0030339
	Teste	97,72%	97,54%	98,45%	98,33%	97,63%	98,04%	98,61%	97,83%	98,00%	98,02%	0,0037723
Naive Bayes	Treino	98,17%	98,08%	98,22%	97,88%	98,20%	98,08%	98,57%	98,31%	97,74%	98,14%	0,0023966
	Teste	98,28%	97,71%	98,55%	98,22%	97,67%	98,08%	98,61%	97,75%	98,17%	98,12%	0,0034857

*Tabela de acurácia*

SER - AWGN												
Decision Tree	SER	2,85%	2,58%	2,29%	2,28%	3,23%	2,96%	2,11%	2,67%	3,50%	2,72%	0,004631
Random Forest	SER	2,26%	2,25%	1,93%	2,06%	2,53%	2,42%	2,22%	2,42%	2,50%	2,29%	0,0020167
Knn	SER	2,07%	2,25%	2,12%	2,06%	2,63%	2,29%	2,06%	2,58%	2,17%	2,25%	0,0021873
SVM Linear	SER	11,83%	17,27%	5,64%	5,53%	7,97%	5,71%	5,67%	5,67%	5,17%	7,83%	0,0412493
SVM kernel	SER	1,61%	2,06%	1,64%	1,64%	2,43%	1,92%	1,50%	2,25%	2,33%	1,93%	0,0035087
MLP	SER	2,28%	2,46%	1,55%	1,67%	2,37%	1,96%	1,39%	2,17%	2,00%	1,98%	0,0037723
Naive Bayes	SER	1,72%	2,29%	1,45%	1,78%	2,33%	1,92%	1,39%	2,25%	1,83%	1,88%	0,0034857

*Tabela da symbol error rate - SER*

✚ **Observação<sub>1</sub>:** Foram levados em consideração a média e o desvio padrão de todos os resultados.

✚ **Observação<sub>2</sub>:** Célula em amarelo é o classificador escolhido.

## ***Indicativo baseado no tempo comparativo para a etapa de teste em uma dada máquina.***

### **Decision Tree**

CPU times: user 65 ms, sys: 0 ns, total: 65 ms  
Wall time: 65.7 ms  
CPU times: user 61.9 ms, sys: 799 µs, total: 62.7 ms  
Wall time: 62.3 ms

### **Random Forest**

CPU times: user 1.17 s, sys: 6.01 ms, total: 1.17 s  
Wall time: 1.23 s  
CPU times: user 1.19 s, sys: 2.17 ms, total: 1.2 s  
Wall time: 1.21 s

### **K-nearest neighbors**

CPU times: user 213 ms, sys: 3.41 ms, total: 217 ms  
Wall time: 477 ms  
CPU times: user 205 ms, sys: 0 ns, total: 205 ms  
Wall time: 212 ms

### **Linear SVM**

CPU times: user 809 ms, sys: 1.67 ms, total: 811 ms  
Wall time: 1.84 s  
CPU times: user 804 ms, sys: 787 µs, total: 805 ms  
Wall time: 1.17 s

### **SVM with RBF Kernel**

CPU times: user 561 ms, sys: 0 ns, total: 561 ms  
Wall time: 561 ms  
CPU times: user 901 ms, sys: 2.38 ms, total: 903 ms  
Wall time: 1.22 s

### **MLP**

CPU times: user 1min 28s, sys: 1min 6s, total: 2min 35s  
Wall time: 1min 19s  
CPU times: user 1min 28s, sys: 1min 8s, total: 2min 36s  
Wall time: 1min 25s

### **Naive Bayes**

CPU times: user 42.9 ms, sys: 579 µs, total: 43.5 ms  
Wall time: 45.2 ms  
CPU times: user 37.2 ms, sys: 1.5 ms, total: 38.7 ms  
Wall time: 39.4 ms