

# FORMAÇÃO INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING

---

MACHINE LEARNING - FUNDAMENTOS  
CLASSIFICAÇÃO PARTE II

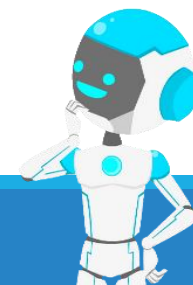
Prof. Fernando Amaral – Todos os Direitos Reservados

# Matriz de Confusão

Idade	Pagou	Classificação
18	Não	Não
46	Sim	Sim
34	Sim	Não
21	Não	Sim
37	Não	Não
...		



		Sim	Não	Classificação
Dados	Sim	1	1	
	Não	1	2	

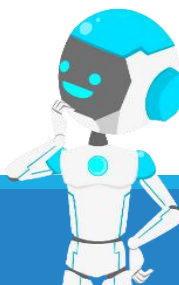


# Matriz de Confusão

---

		Sim	Não	Classificação
Dados	Sim	1	1	
	Não	1	2	

Verdadeiros Positivos	Falsos Negativos
Falsos Positivos	Verdadeiros Negativos



# Generalização Versus Super Ajuste

---

- ❖ O objetivo de todo classificador é criar modelos genéricos
- ❖ O modelo super ajustado funciona bem com dados de treino, mas tem o desempenho pobre em dados de teste ou de produção.



Genérico



Super Ajustado



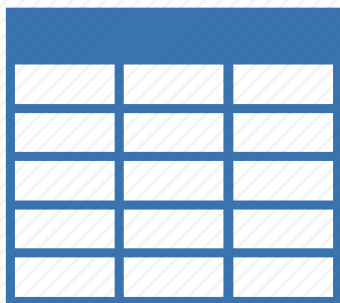
# Causas de Super Ajuste

---

- Dados não representativos
- Dados não significativos (poucos)
- Forma de treinamento
- Classe rara
- Modelo incorreto



# Problema da Classe Rara







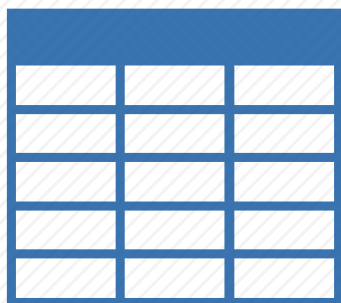
1 milhões de transações  
5% fraudulentas



Amostragem Aleatória  
sem reposição: 1%



10.000  
9.900 - Legítimas  
100 - Fraudulentas







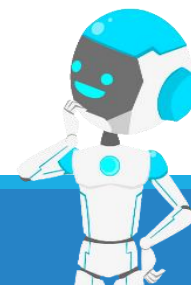
1 milhões de transações  
5% fraudulentas



Amostragem Aleatória  
sem reposição:  
0,5% - Legítimas  
0,5% - Fraudulentas



10.000  
5.000 - Legítimas  
5.000 - Fraudulentas



## Problemas de Atributos Desconhecidos

---

- No treino: regiões “Sul”, “Sudeste”, “Centro-Oeste” e “Norte”
- Na produção: região “Nordeste”



# Métricas

Originais	Treino		
	Aprovado		Reprovado
	Aprovado	1004	208
	Reprovado	202	294

Métrica	Fórmula	Descrição	Cálculo
Acertos	$(VP+VN)/Total$	Total de Acertos	75,99
Erros	$(FP+FN)/Total$	Total de Erros	24,00
Precisão	$VP/(VP+FP)$	Quantos registros de fato são positivos	83,25
Lembrança ou Positivos Verdadeiros	$VP/(VP+FN)$	Positivos corretamente previstos	82,83
Negativos Verdadeiros	$VN/(VN+FP)$	Total de Negativos Falsos	59,27
Positivos Falsos	$FP/(VN+FP)$	Total de Positivos Falsos	40,72
Negativos Falsos	$FN/(VP+FN)$	Total de Negativos Falsos	17,16

