



CLASSIFICADORES MULTICLASSE



Allan Dieguez

AI Researcher | Data Scientist

Director, Data Science

BAIN & COMPANY 

LinkedIn: *@allandieguez*

E-Mail: *allandieguez@gmail.com*



AGENDA

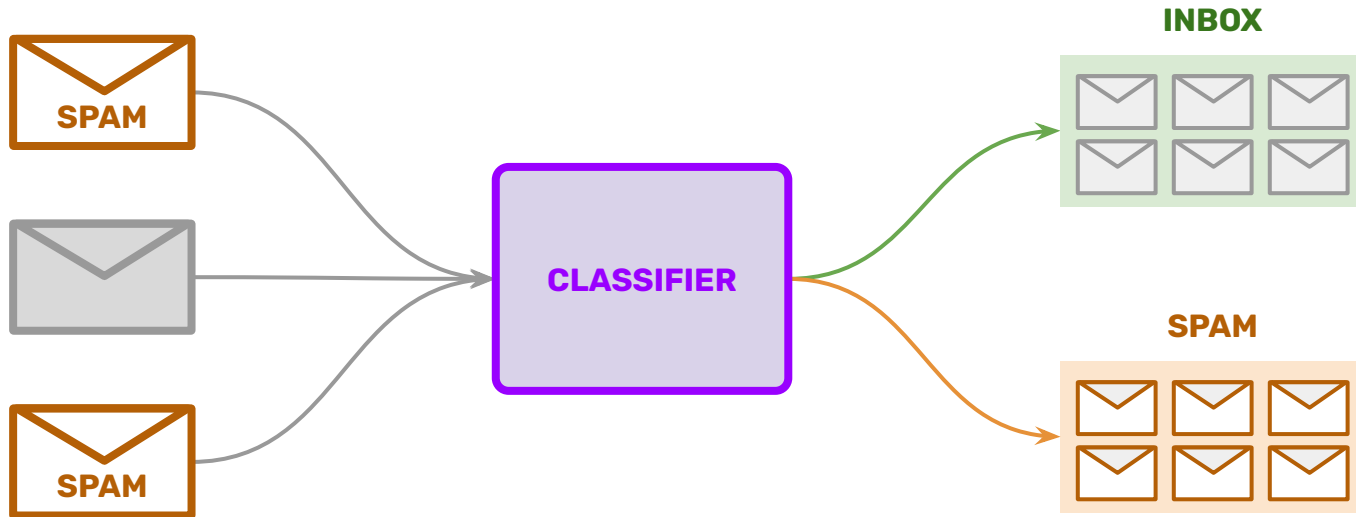
- **Problemas Multiclasse**
Complexidade com mais classes
- **Combinação de Classificadores**
Soluções com modelos binários
- **Arquiteturas Multiclasse**
Suporte nativo a multiclasse
- **Métricas de Avaliação**
O que muda com mais classes



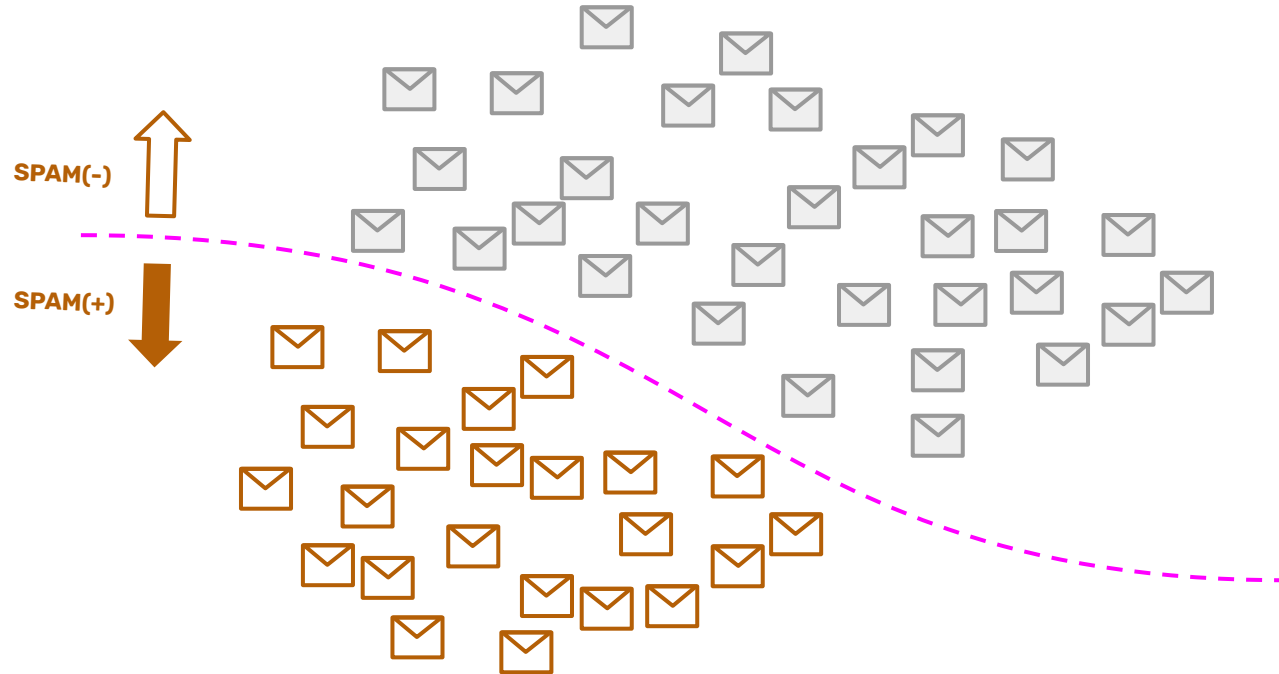
PROBLEMAS MULTICLASSE

Complexidade com mais classes

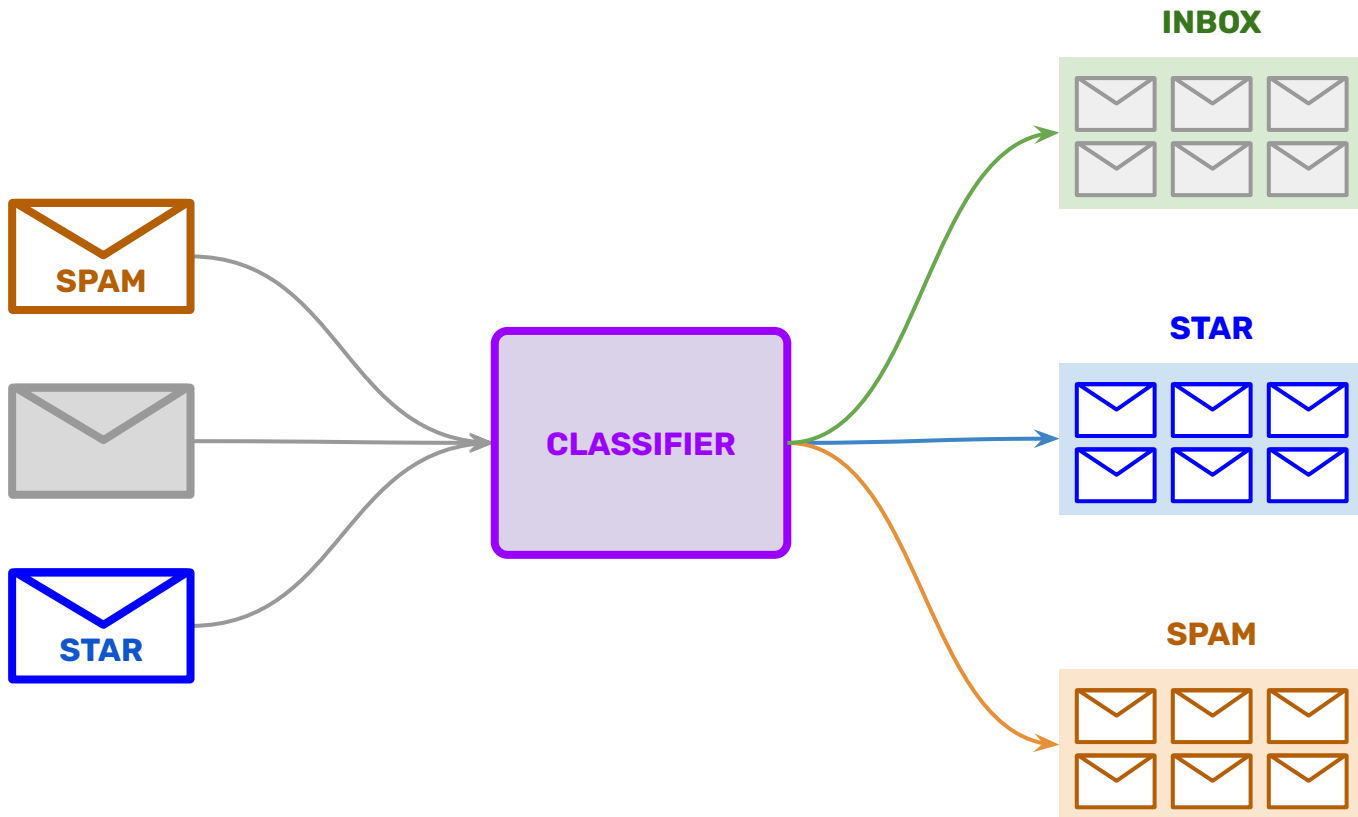
A CLASSIFICAÇÃO BINÁRIA



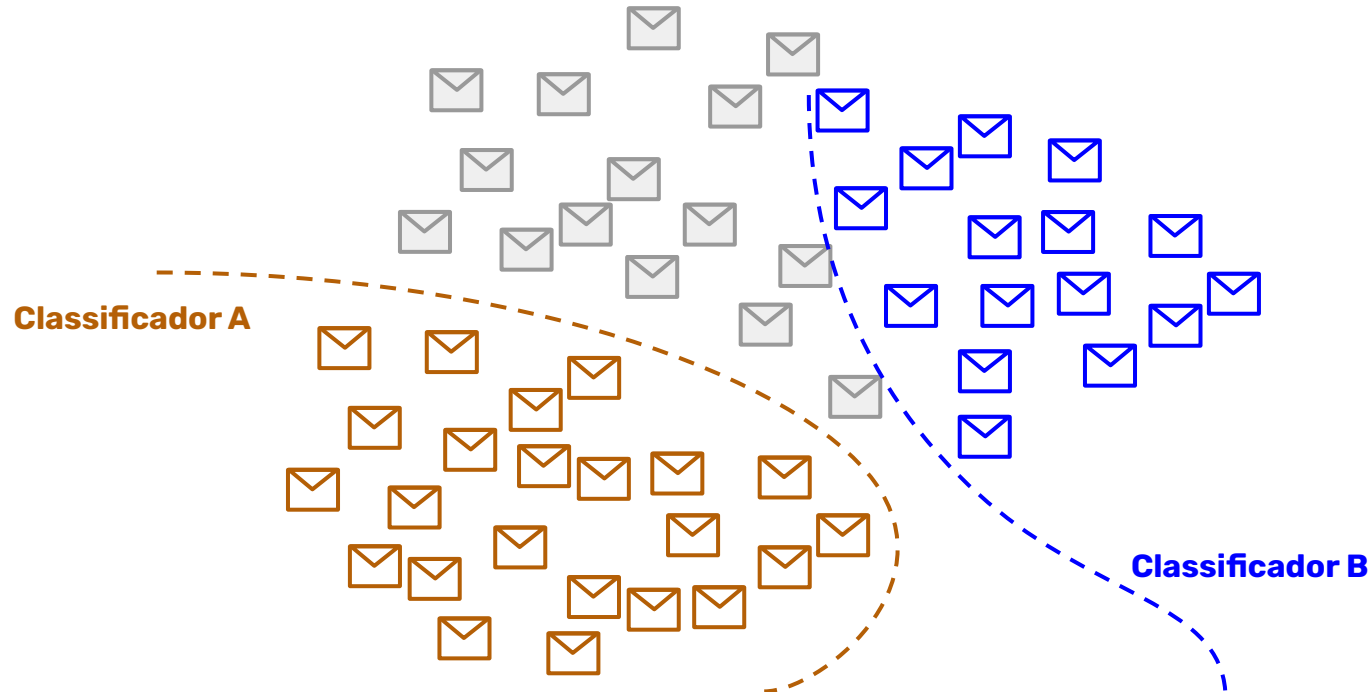
HIPERPLANO DE DECISÃO



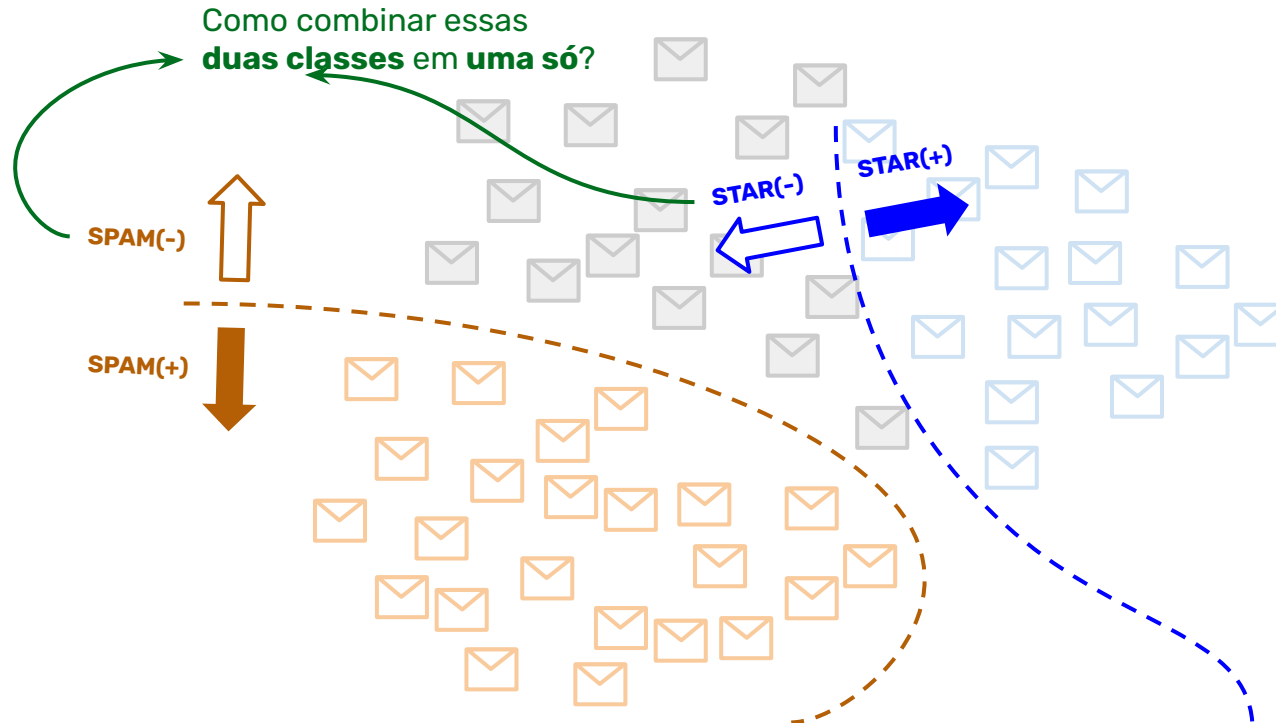
COMO SOLUCIONAR?



MAIS COMPLEXIDADE COM MULTICLASSE



MODELOS COMBINADOS





COMBINAÇÃO DE CLASSIFICADORES

Soluções com modelos binários

	Classe 1	Classe 2	Classe 3	...	Classe N
Modelo 1	+	-	-	...	-
Modelo 2	-	+	-	...	-
...			...		
Modelo N	-	-	-	...	+

TÉCNICA ONE VS ALL

São treinados **N** modelos, **um para cada classe**, particionando em cada um o domínio em **classe** e não **classe**.

A decisão final é computada como **scores de pertinência** a cada classe.

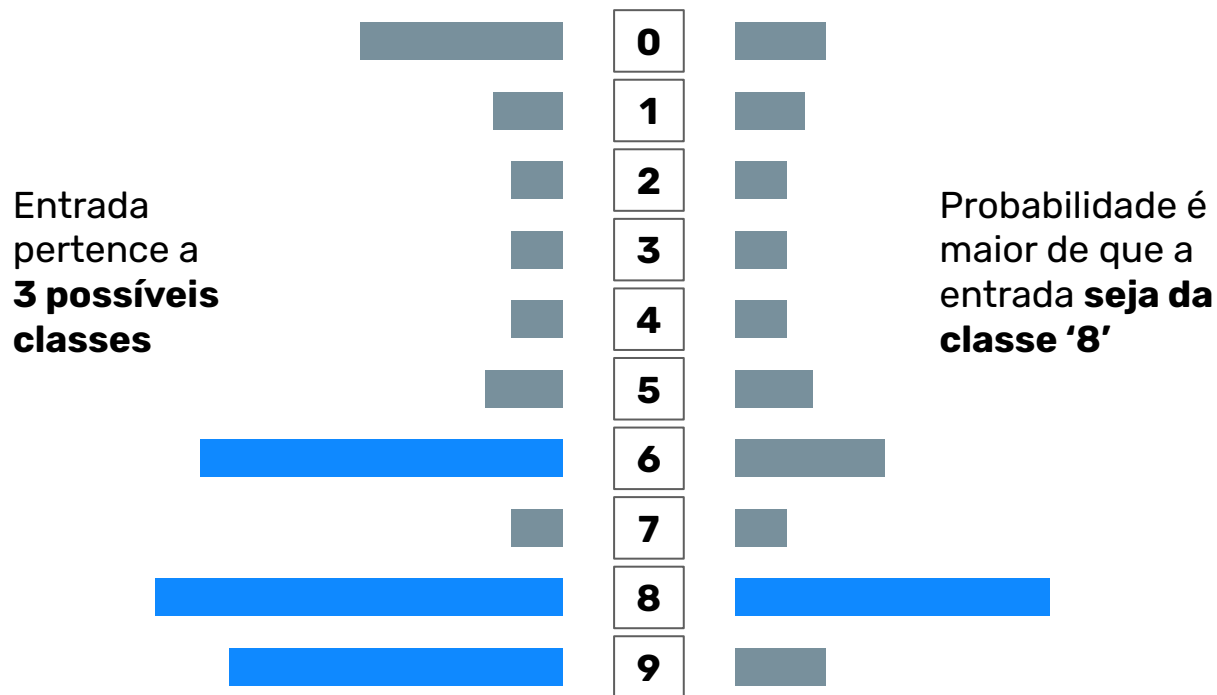
	Classe 1	Classe 2	Classe 3	...	Classe N
Modelo 1	+	-	X	...	X
Modelo 2	X	+	-	...	X
...			...		
Modelo K	X	X	-	...	+

TÉCNICA ONE VS ONE

São treinados **K** modelos, um para cada **duas classes**, onde o valor de **K** é igual à **combinatória $C(N, 2)$** .

A decisão final pode ser computada a partir de **votação simples** ou usando uma heurística sobre os *scores* de cada saída.

PERTINÊNCIA VS PROBABILIDADE

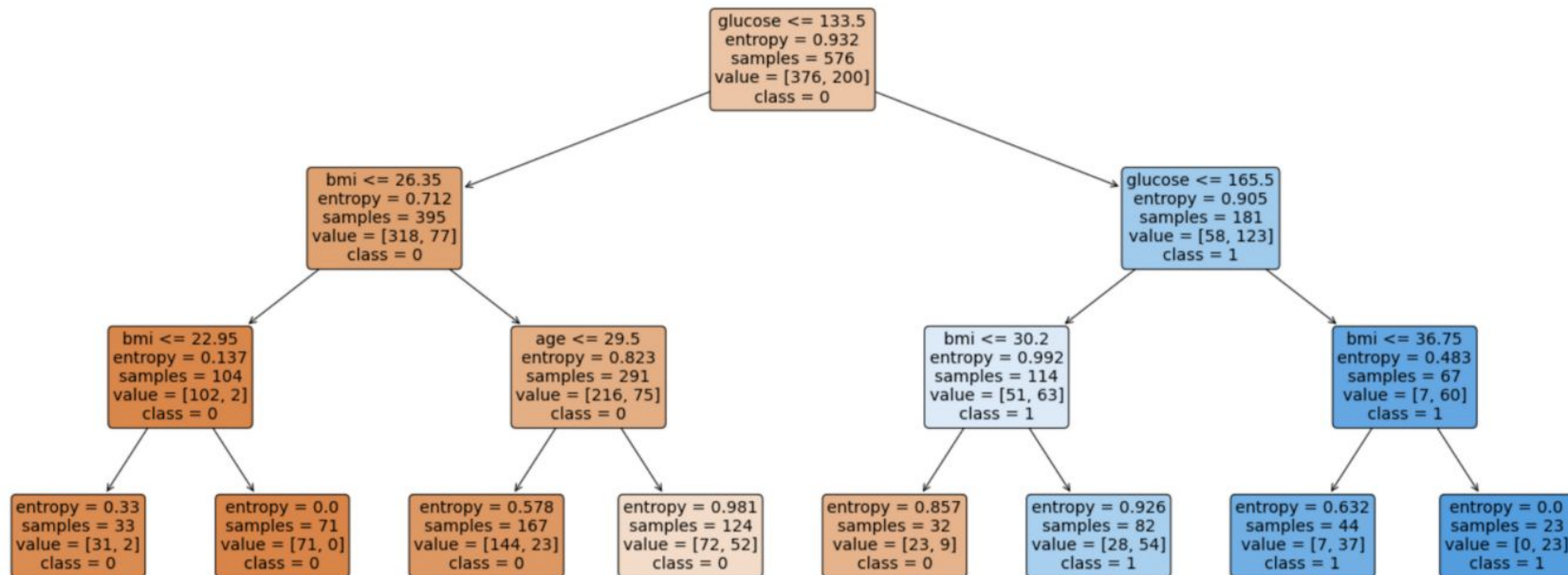


A vertical bar with a gradient from green at the top to blue at the bottom.

ARQUITETURAS MULTICLASSE

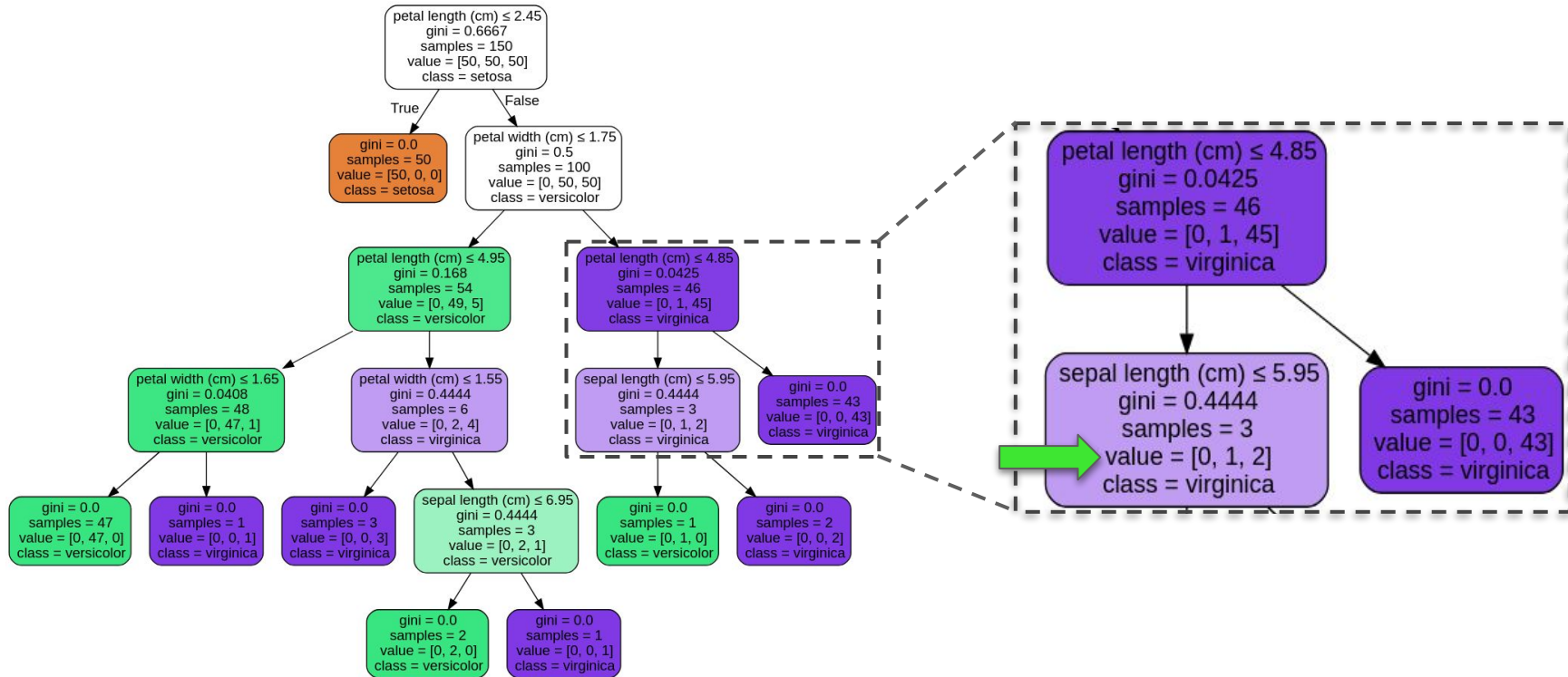
Suporte nativo a multiclasse

ÁRVORE DE DECISÃO BINÁRIA

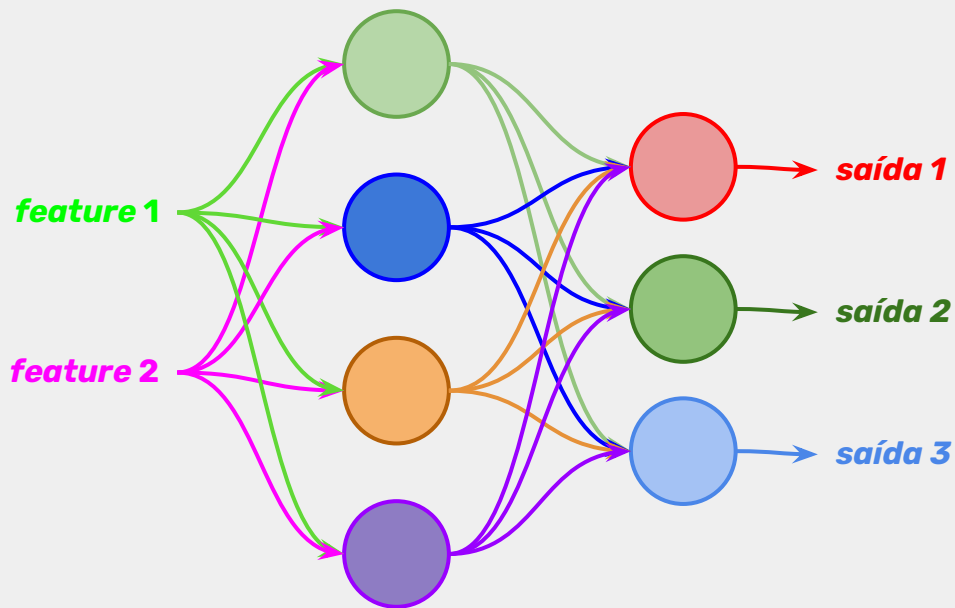


Fonte: [Artigo "Binary Classification using Decision-Tree Model" do Medium](#)

ÁRVORE DE DECISÃO MULTICLASSE



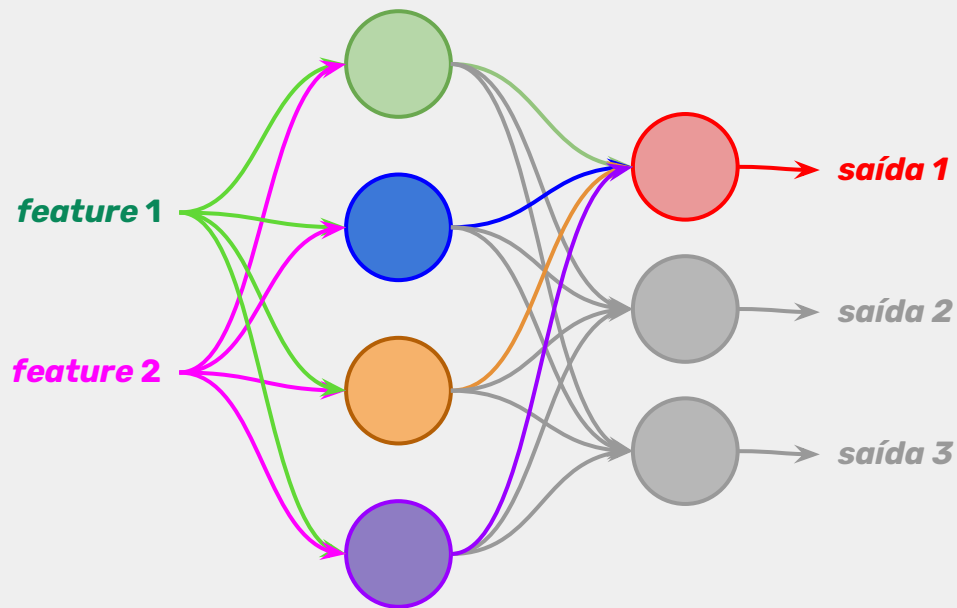
Fonte: [tutorial do Scikit-Learn sobre classificação sobre o dataset Iris](#)



MULTI LAYER PERCEPTRON

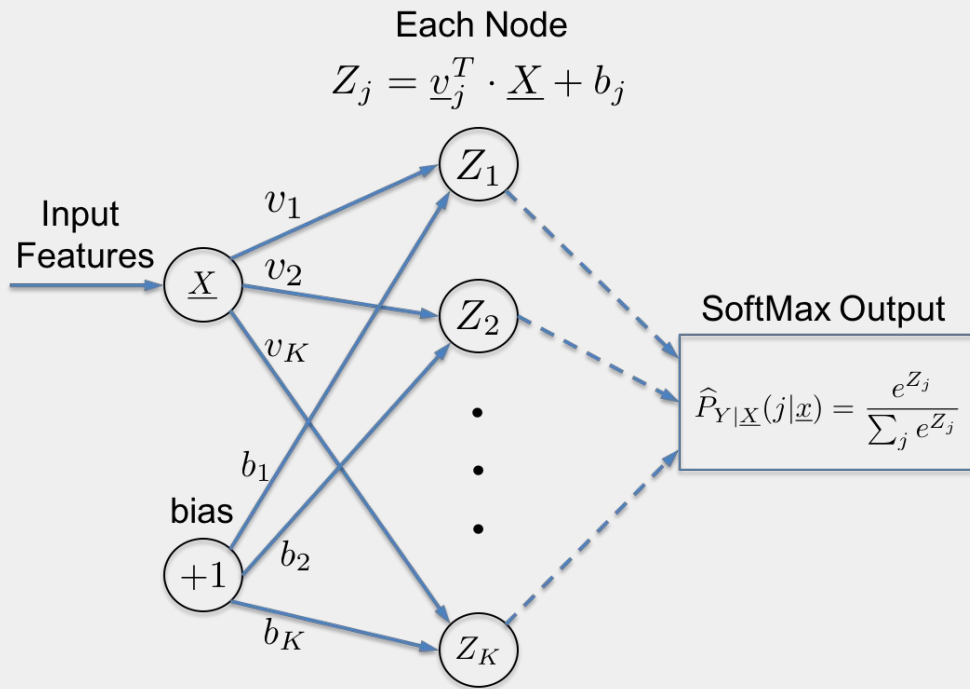
Estrutura **hiperconectada** de neurônios artificiais (modelos matemáticos baseados na biologia).

Cada neurônio equivale a uma **regressão logística**.



ESTRUTURA COMPARTILHADA

A decisão em **cada neurônio de saída** é alimentada pela **mesma estrutura** que passa informações aos outros.



FUNÇÃO SOFTMAX

A função Softmax força que **todas as saídas** do modelo em conjunto emulem uma função de **densidade de probabilidade** discretizada.



MÉTRICAS DE AVALIAÇÃO

O que muda com mais classes

		Predicted Label	
		False	True
True Label	False	Verdadeiro Negativo	Falso Positivo
	True	Falso Negativo	Verdadeiro Positivo

Matriz de Confusão Binária

Métrica

Racional

Fórmula

Acurácia
(Accuracy)

Elementos **corretamente classificados** como **positivos** ou **negativos**

$$\frac{VP + VN}{Total}$$

Precisão
(Precision)

Dentre as **previsões positivas**, quantas são os elementos **rotulados positivos**?

$$\frac{VP}{VP + FP}$$

Sensibilidad e
(Recall)

Dentre os elementos **rotulados positivos**, quantas são as **previsões positivas**?

$$\frac{VP}{VP + FN}$$

MÉTRICA ÚNICA DE PERFORMANCE

$$F_{\beta} = (1+\beta^2) \times \frac{\textit{Precision} \times \textit{Recall}}{\beta^2 \times \textit{Precision} + \textit{Recall}}$$

Média harmônica ponderada entre as medidas de **precisão** e **sensibilidade**.

$$F_1 = 2 \times \frac{\textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

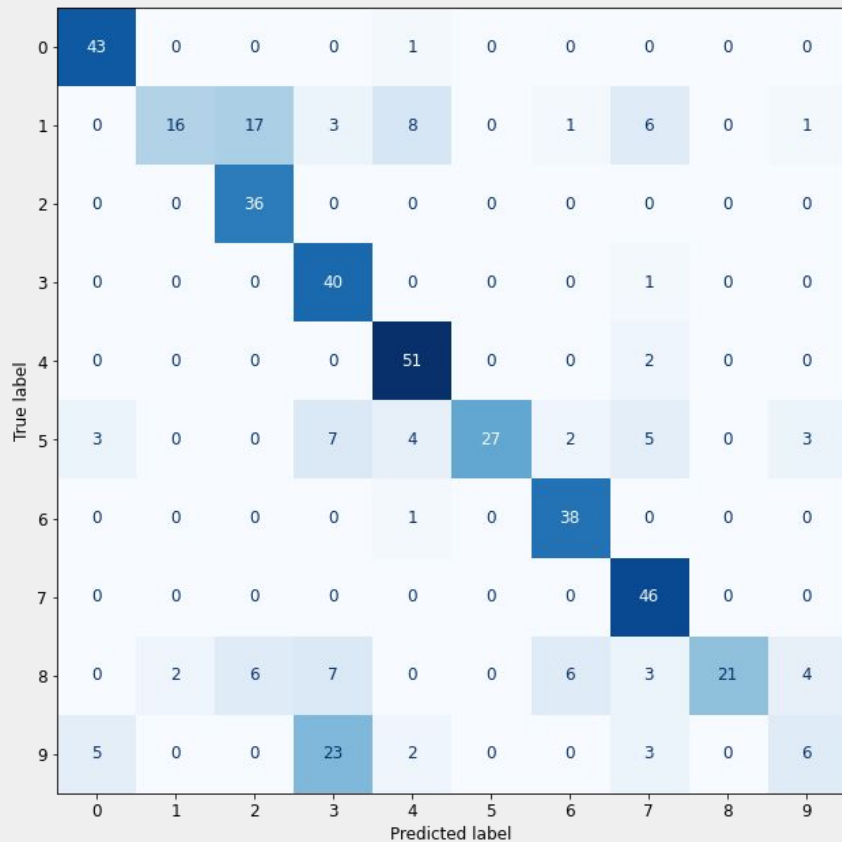
Mais utilizada por equilibrar **ambas as medidas** ao mesmo tempo.

$$F_2 = 5 \times \frac{\textit{Precision} \times \textit{Recall}}{4 \times \textit{Precision} + \textit{Recall}}$$

Utilizada quando a maior ênfase da métrica deve ser dada à **precisão** (ou **precision**)

$$F_{0.5} = 1.25 \times \frac{\textit{Precision} \times \textit{Recall}}{0.25 \times \textit{Precision} + \textit{Recall}}$$

Utilizada quando a maior ênfase da métrica deve ser dada à **sensibilidade** (ou **recall**)



MATRIZ DE CONFUSÃO MULTICLASSE

Como calcular

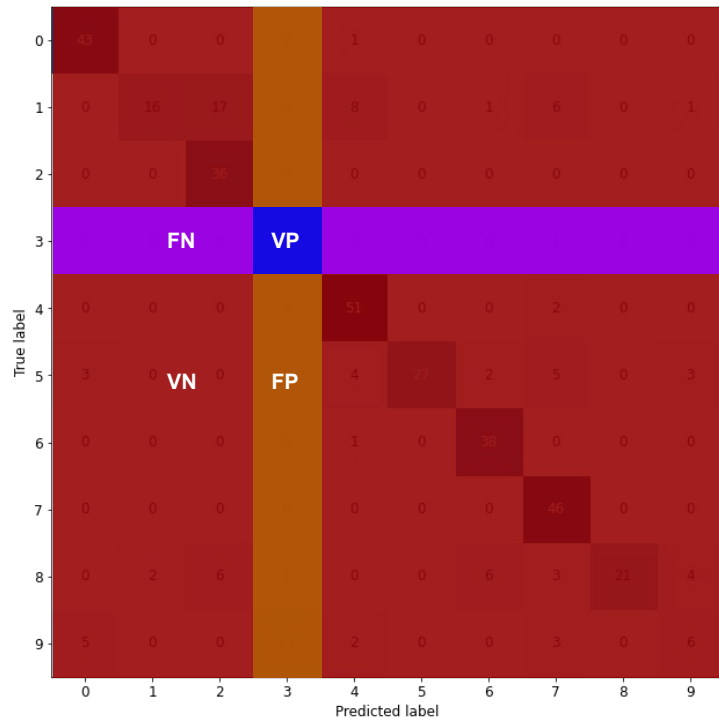
Falso Positivo,

Falso Negativo,

**Verdadeiro Positivo e Verdadeiro
Negativo**

na Matriz de Confusão Multiclasse?

MÉTRICAS DE CONFUSÃO MULTICLASSE



	Accuracy	Precision	Recall	F1-Score
Classe 0	0.98	0.84	0.98	0.91
Classe 1	0.92	0.89	0.31	0.46
Classe 2	0.95	0.61	1.00	0.76
Classe 3	0.91	0.50	0.98	0.66
Classe 4	0.96	0.76	0.96	0.85
Classe 5	0.95	1.00	0.53	0.69
Classe 6	0.98	0.81	0.97	0.88
Classe 7	0.96	0.70	1.00	0.82
Classe 8	0.94	1.00	0.43	0.60
Classe 9	0.91	0.43	0.15	0.23

CALCULAR UMA ÚNICA MÉTRICA

Método	Racional	Efeito
<i>'micro'</i>	Calcula globalmente os valores de FP , FN , VP e VN para calcular as métricas	Racional mais próximo do original binário
<i>'macro'</i>	Calcula as métricas por classe , individualmente, e faz a média simples	Ignora o desbalanceamento ; <i>F1Score</i> pode sair da faixa [precision, recall]
<i>'weighted'</i>	Calcula a média ponderada por suporte das métricas por classe	Racional simples que considera o desbalanceamento das classes

Fonte do método: [sklearn.metrics.precision_recall_fscore_support](#)

MÉTRICA ÚNICA POR MODELO

PER CLASS	Precision	Recall	F1Score	Suporte
Classe 0	0.84	0.98	0.91	44
Classe 1	0.89	0.31	0.46	52
Classe 2	0.61	1.00	0.76	36
Classe 3	0.50	0.98	0.66	41
Classe 4	0.76	0.96	0.85	53
Classe 5	1.00	0.53	0.69	51
Classe 6	0.81	0.97	0.88	39
Classe 7	0.70	1.00	0.82	46
Classe 8	1.00	0.43	0.60	49
Classe 9	0.43	0.15	0.23	39

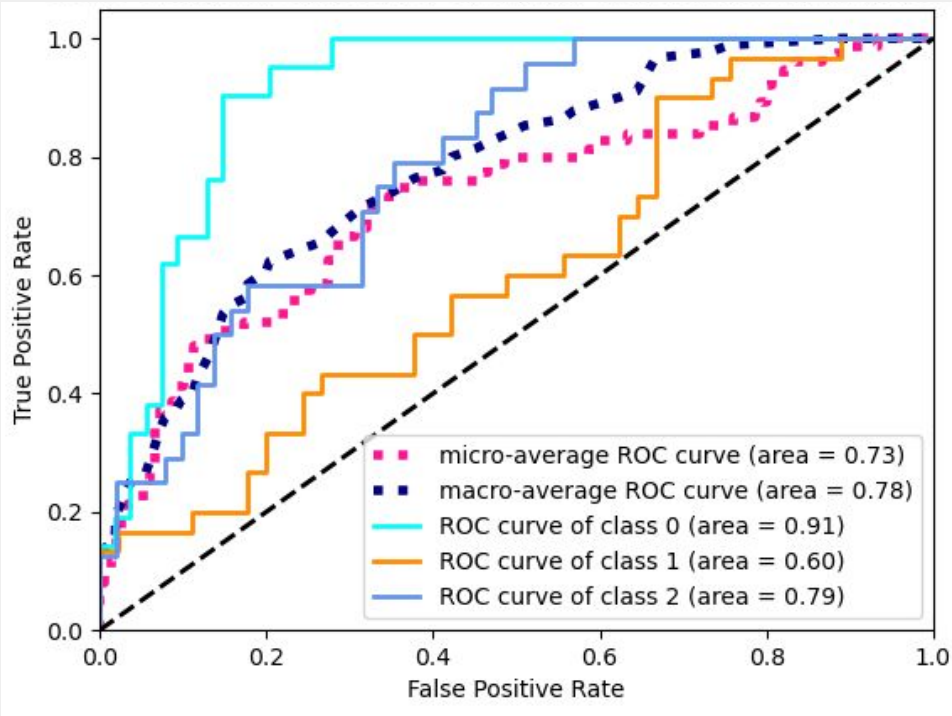
AVERAGE	Precision	Recall	F1Score
'micro'	0.77	0.77	0.77
'macro'	0.80	0.78	0.75
'weighted'	0.81	0.77	0.75

ROC EM PROBLEMAS MULTICLASSE

No cálculo do ROC Multiclasse, o **mesmo racional** descrito para métricas de confusão se aplica **sem restrições**.

Formas de avaliar **ROC Multiclasse**:

- ROC **por classe**;
- Agregando por **Médias**
 - 'micro-average'
 - 'macro-average'
 - 'weighted-average'




Fonte do método: [tutorial do Scikit-Learn, aplicação de ROC Multiclasse](#)

T

A vertical bar with a gradient from green at the top to blue at the bottom.


RESUMO DA AULA

TAKEAWAY #1




Problemas multiclasse trazem mais complexidade quando comparados a problemas de classificação binária

TAKEAWAY #2



Podemos combinar classificadores binários para resolver problemas multiclasse, combinando as saídas dos classificadores.

TAKEAWAY #3



Temos arquiteturas que já nascem multiclasse, tendo suporte nativo para este tipo de problema.

TAKEAWAY #4

Temos métricas para fazer a avaliação dos classificadores. Além disso, métricas voltadas para problemas binários podem ser utilizadas para problemas multiclasse.

