

AVALIANDO O DESEMPENHO DO RANDOM FOREST E REDES NEURAIS EM PROBLEMAS DE CLASSIFICAÇÃO BINÁRIA

VINÍCIUS BARBOSA FARIA – 7312

BIANCA PANACHO FERREIRA – 7546

PEDRO HENRIQUE CAMPOS MOREIRA – 8745



SUMÁRIO

01 INTRODUÇÃO

02 MATERIAIS E MÉTODOS

03 RESULTADOS

04 CONCLUSÃO



INTRODUÇÃO

O Random Forest é uma combinação de classificadores em estrutura de árvore:

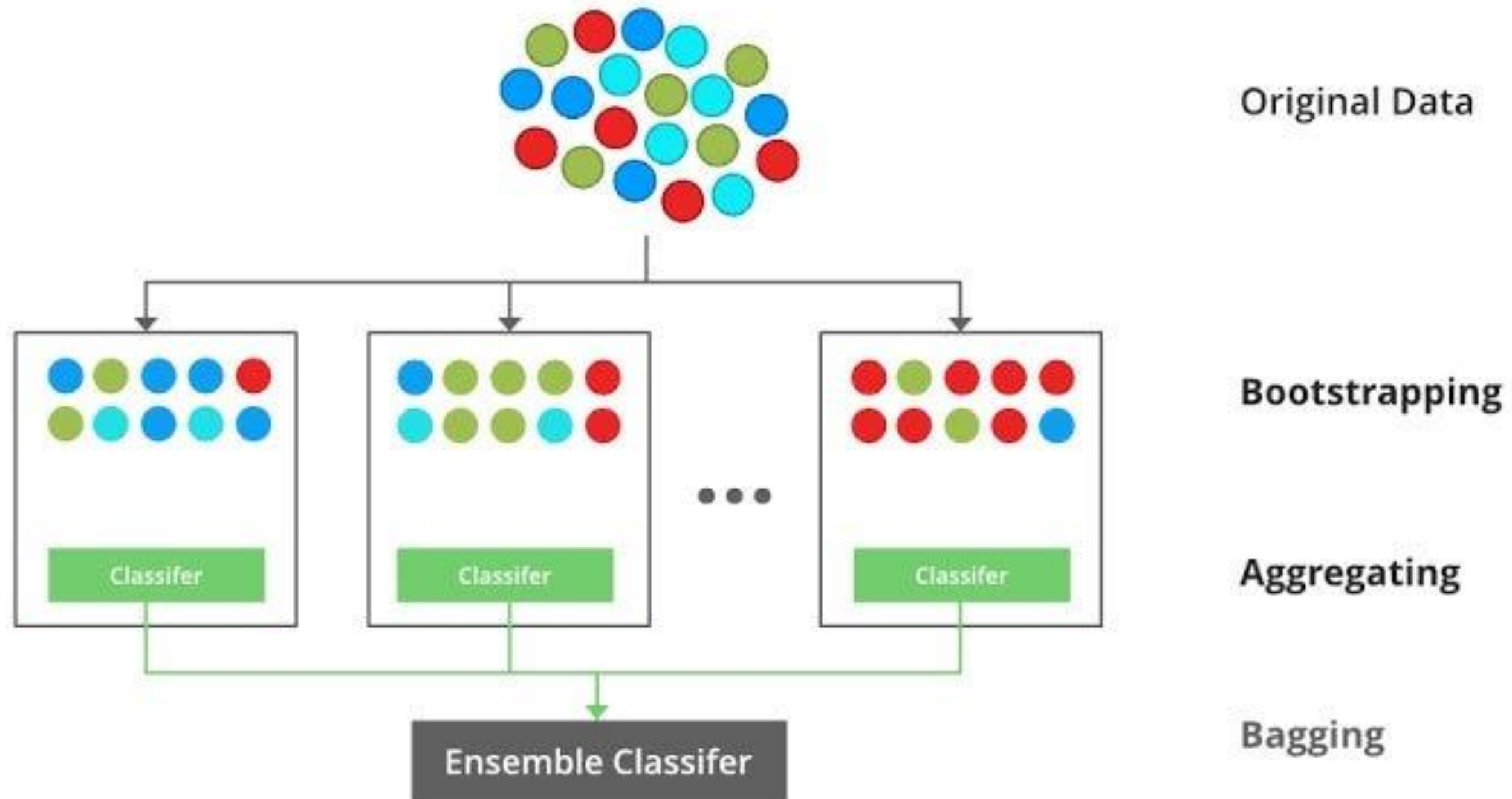
- Classificação, previsão e regressão.

A Rede Neural - Perceptron:

- Classificador linear (binário);
- Aprendizagem supervisionada;
- Várias entradas, x_1 , x_2 , x_N e produz uma única saída binária.

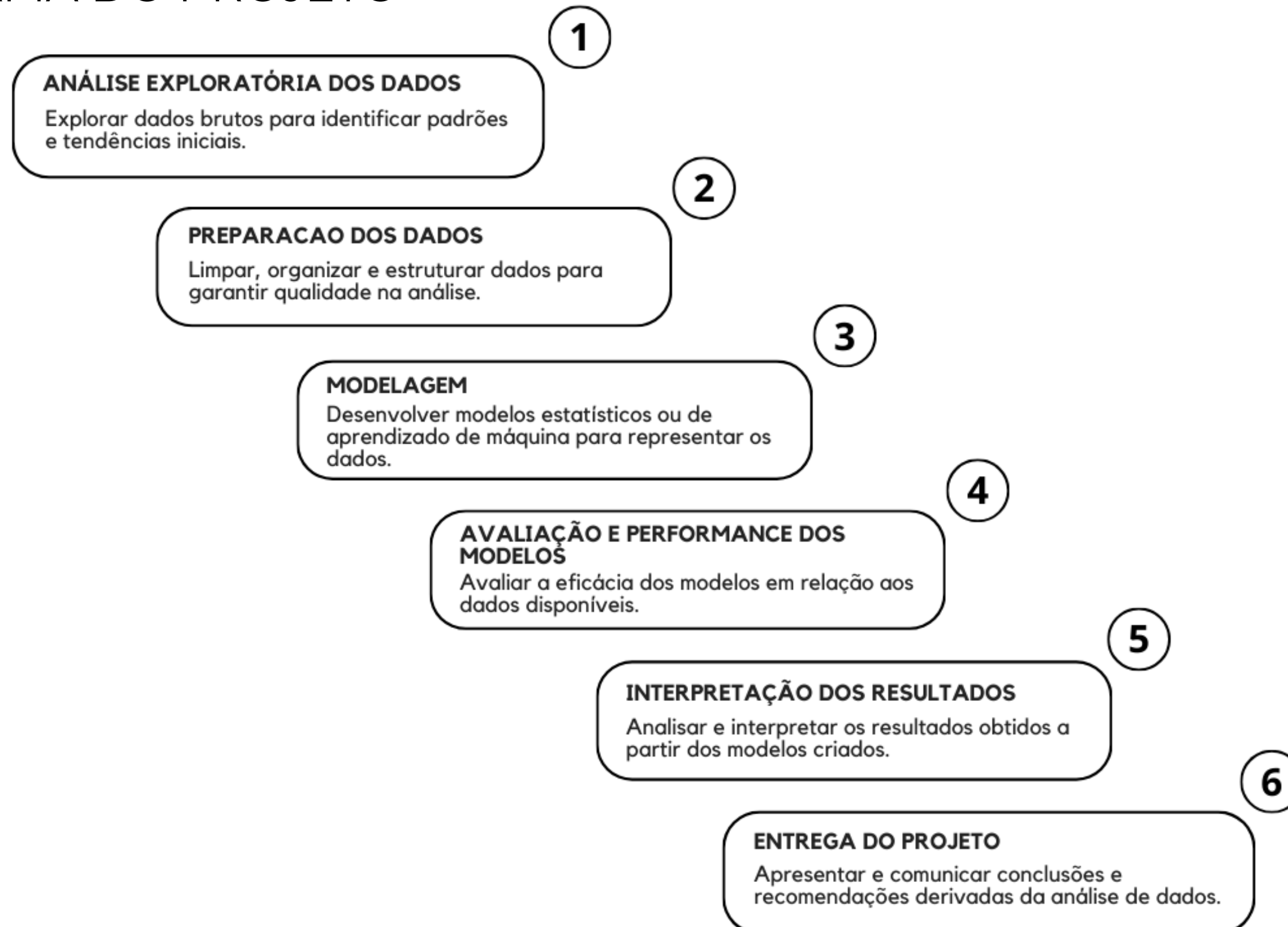
MATERIAL E MÉTODOS

FUNCIONAMENTO DO RANDOM FOREST



MATERIAL E MÉTODOS

FLUXOGRAMA DO PROJETO



MATERIAL E MÉTODOS

MÉTRICAS UTILIZADAS

$$Acuracia = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$Precisao = \frac{VP}{VP + FP} \quad (2)$$

$$Recall = \frac{VP}{VP + FN} \quad (3)$$

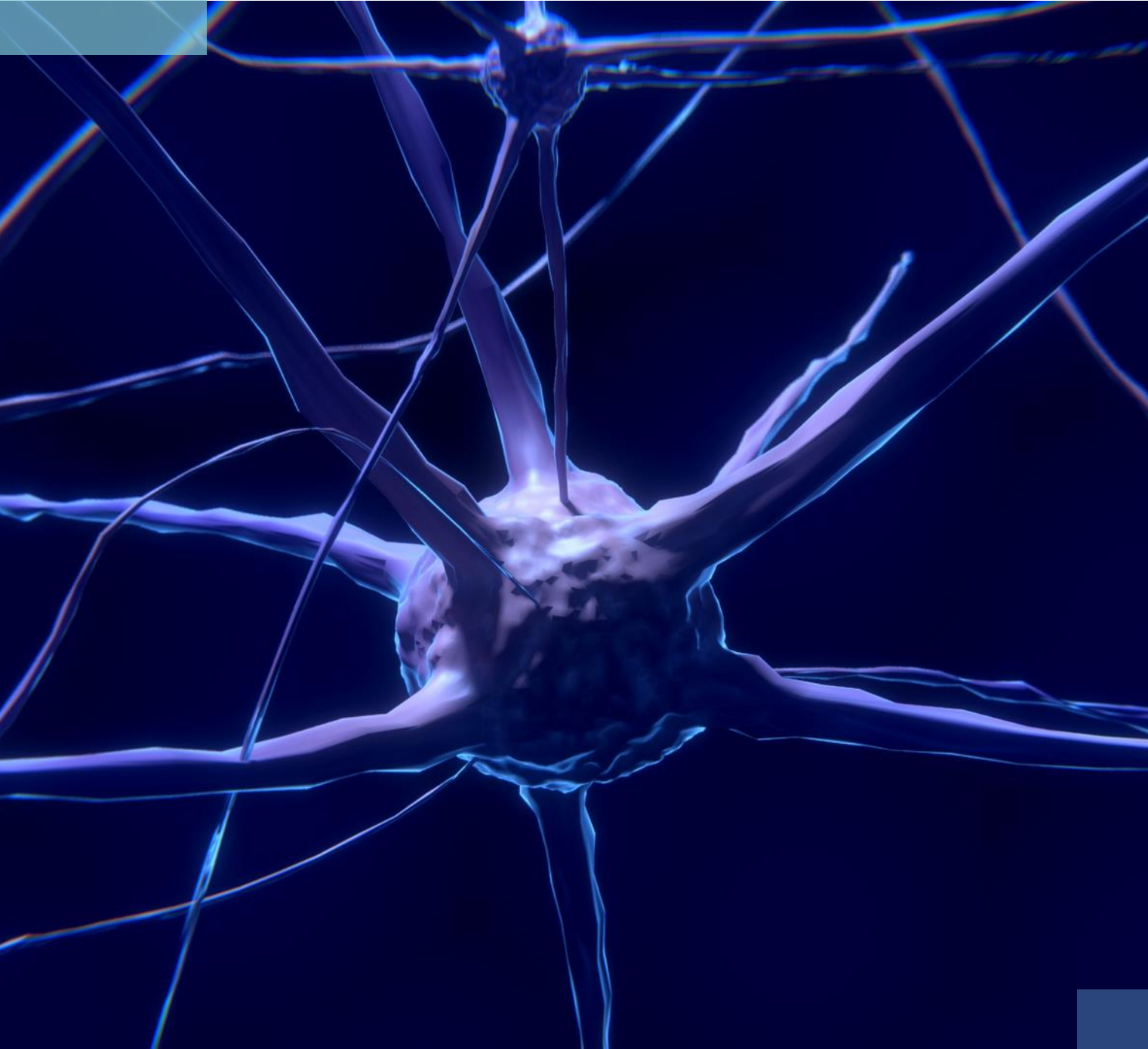
$$F1 - Score = 2 \times \frac{Precisao \times Recall}{Precisao + Recall} \quad (4)$$

$$SP = \text{Número De Instâncias Da Classe} \quad (5)$$

$$Macro = \frac{Metrica(0) + Metrica(1)}{2} \quad (6)$$

$$Ponderada = \frac{SP(0) \times Metrica(1) + SP(1) \times Metrica(0)}{SPTotal} \quad (7)$$

MATERIAL E MÉTODOS



REDE PERCEPTRON MLP (MULTICAMADAS)

- 1 CAMADA
- 2 CAMADAS INTERMEDIÁRIAS
- 3 ÚLTIMA CAMADA

RESULTADOS

O modelo Random Forest apresentou uma acurácia de 74.04%:

Classe 0

Precisão: 74%

Revocação (Recall):
80%

F1-Score: 77%

Suporte: 56 instâncias

Classe 1

Precisão: 74%

Revocação (Recall):
67%

F1-Score: 70%

Suporte: 48 instâncias

RESULTADOS

O modelo MLP apresentou uma acurácia de 72.00%:

Relatório

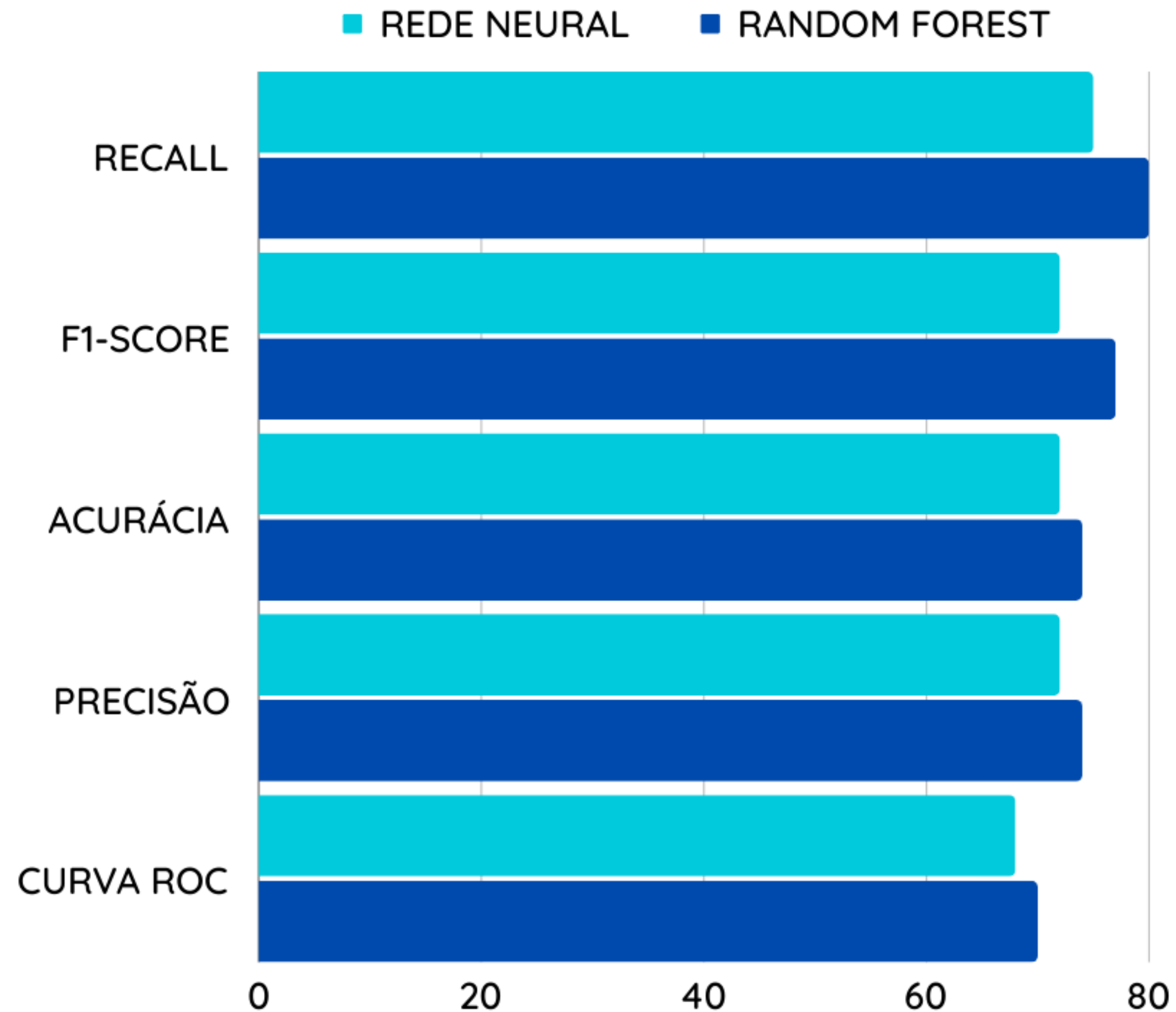
Precisão: 72%

Revocação (Recall): 80%

F1-Score: 74%

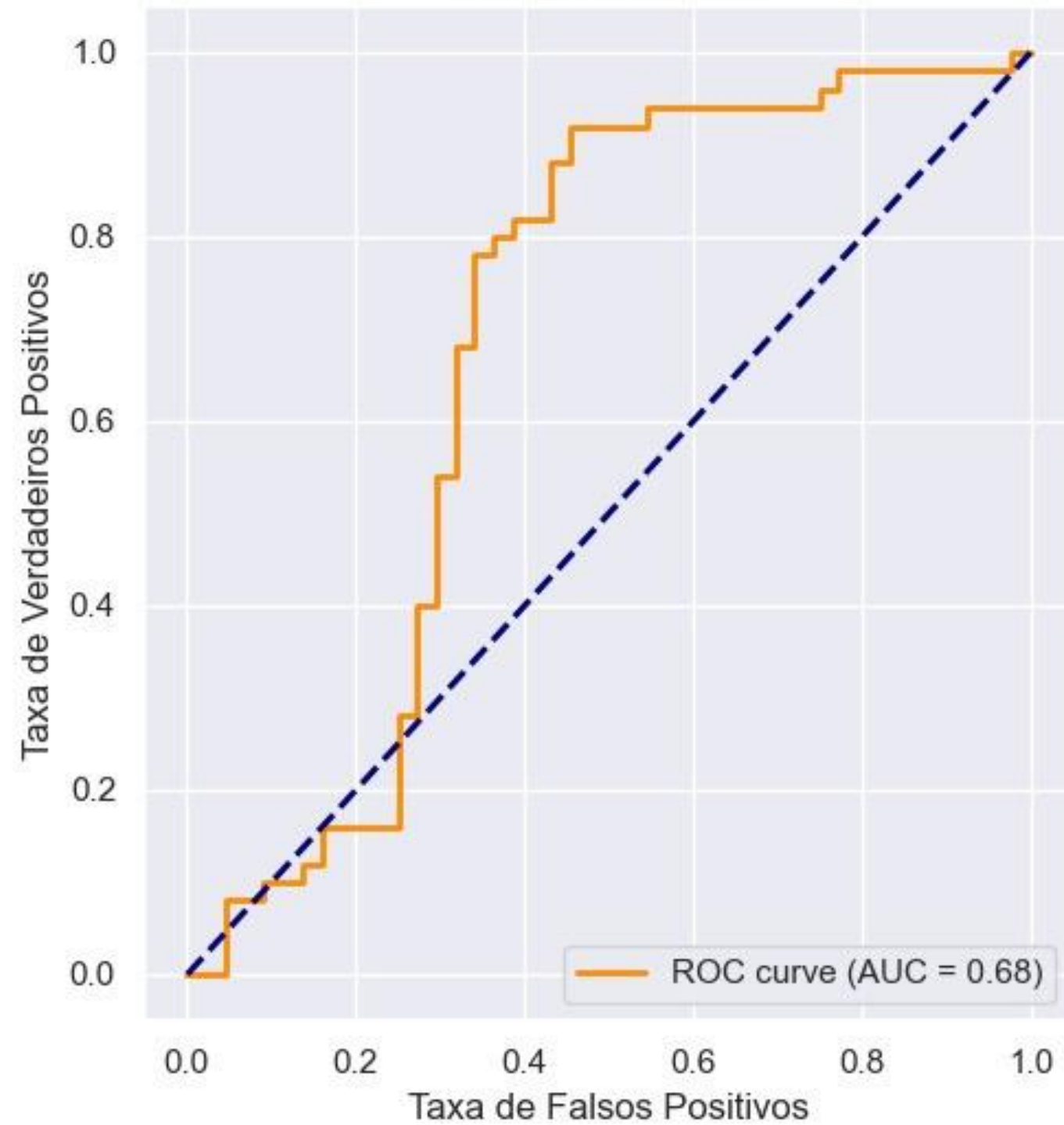
Suporte: 50 instâncias

MÉTRICAS PERCEPTRON MLP VS. RANDOM FOREST

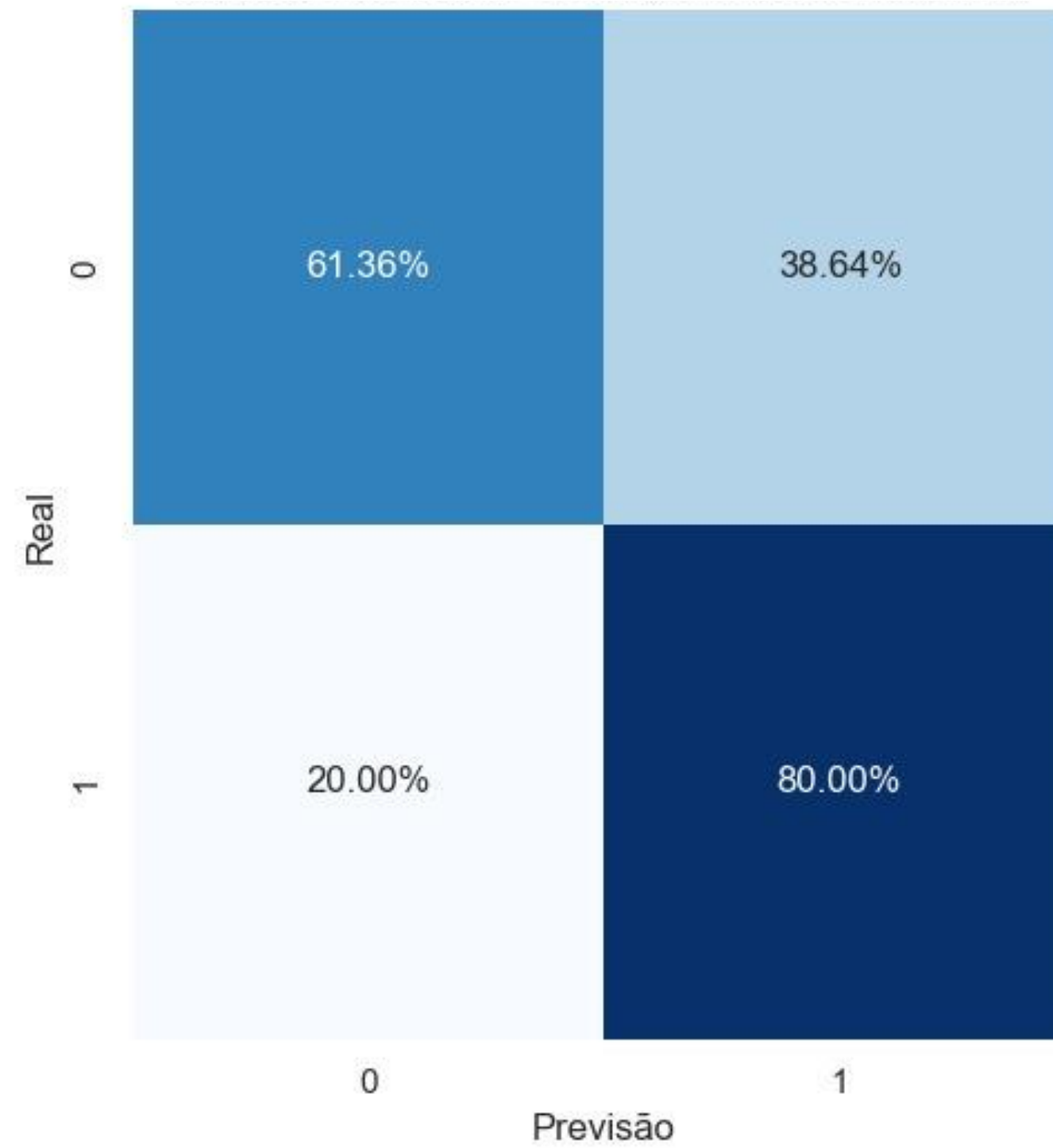


PERCEPTRON MULTICAMADAS

Curva ROC (MLP)

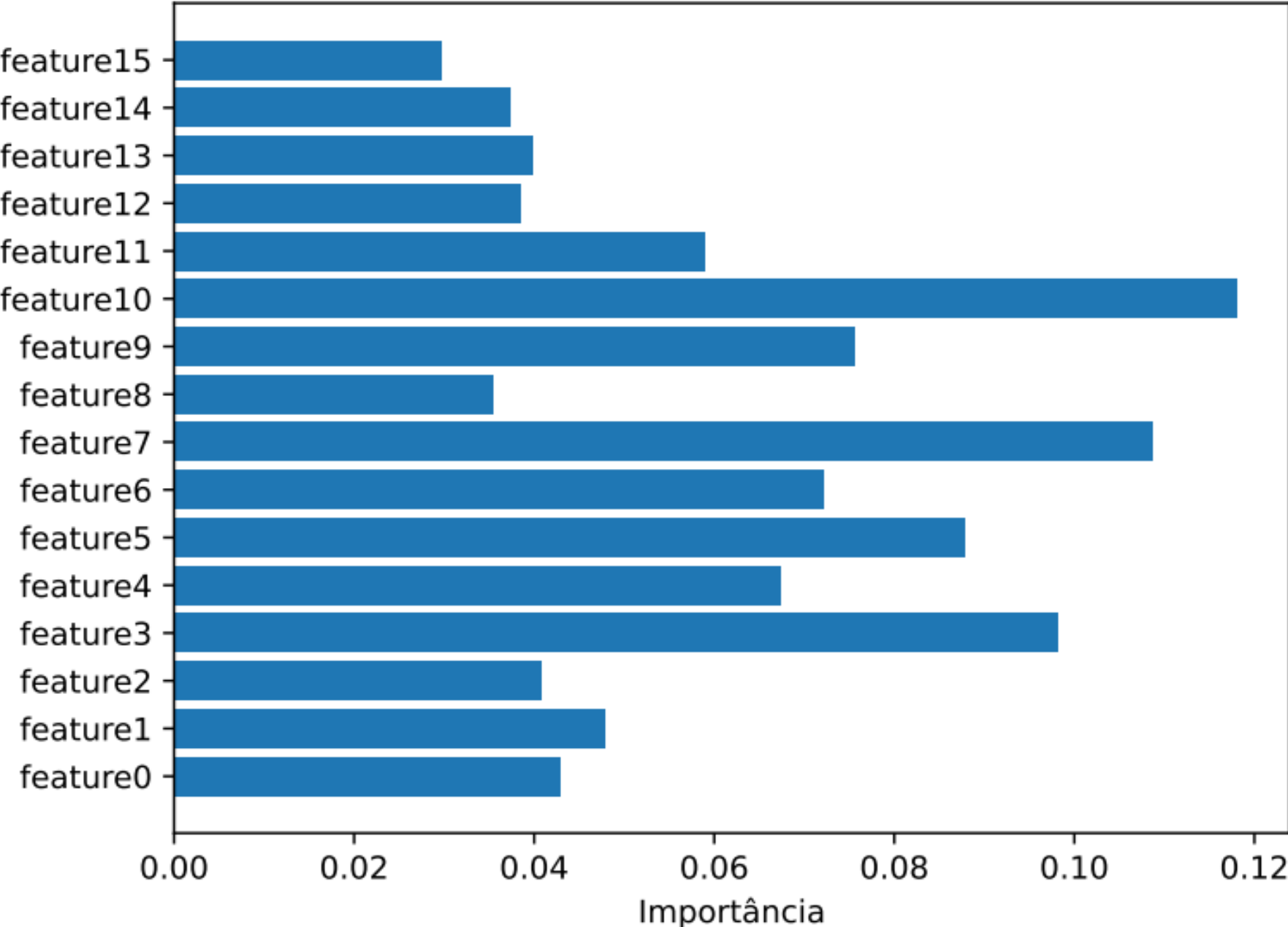


Matriz de Confusão (MLP)

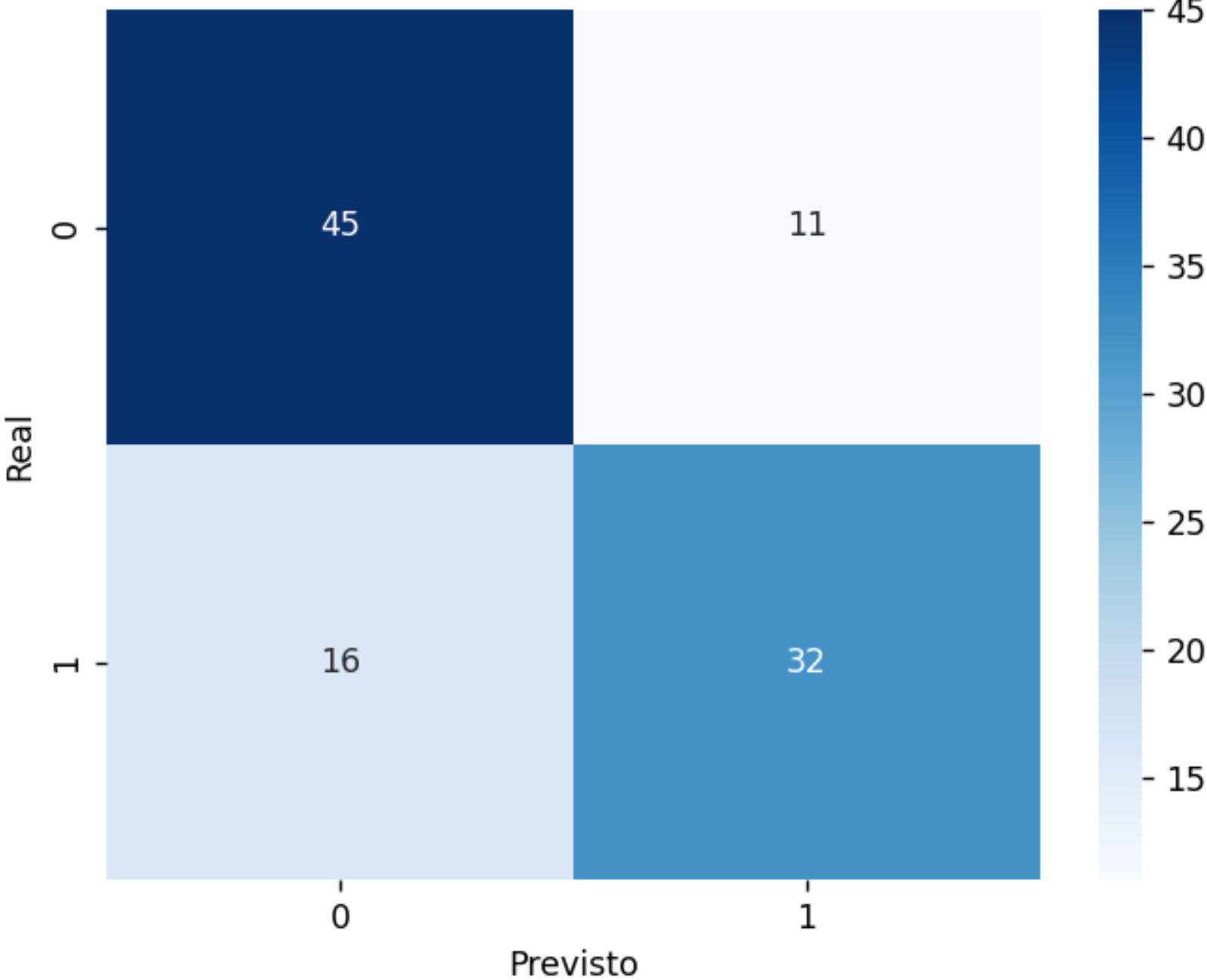


RANDOM FOREST

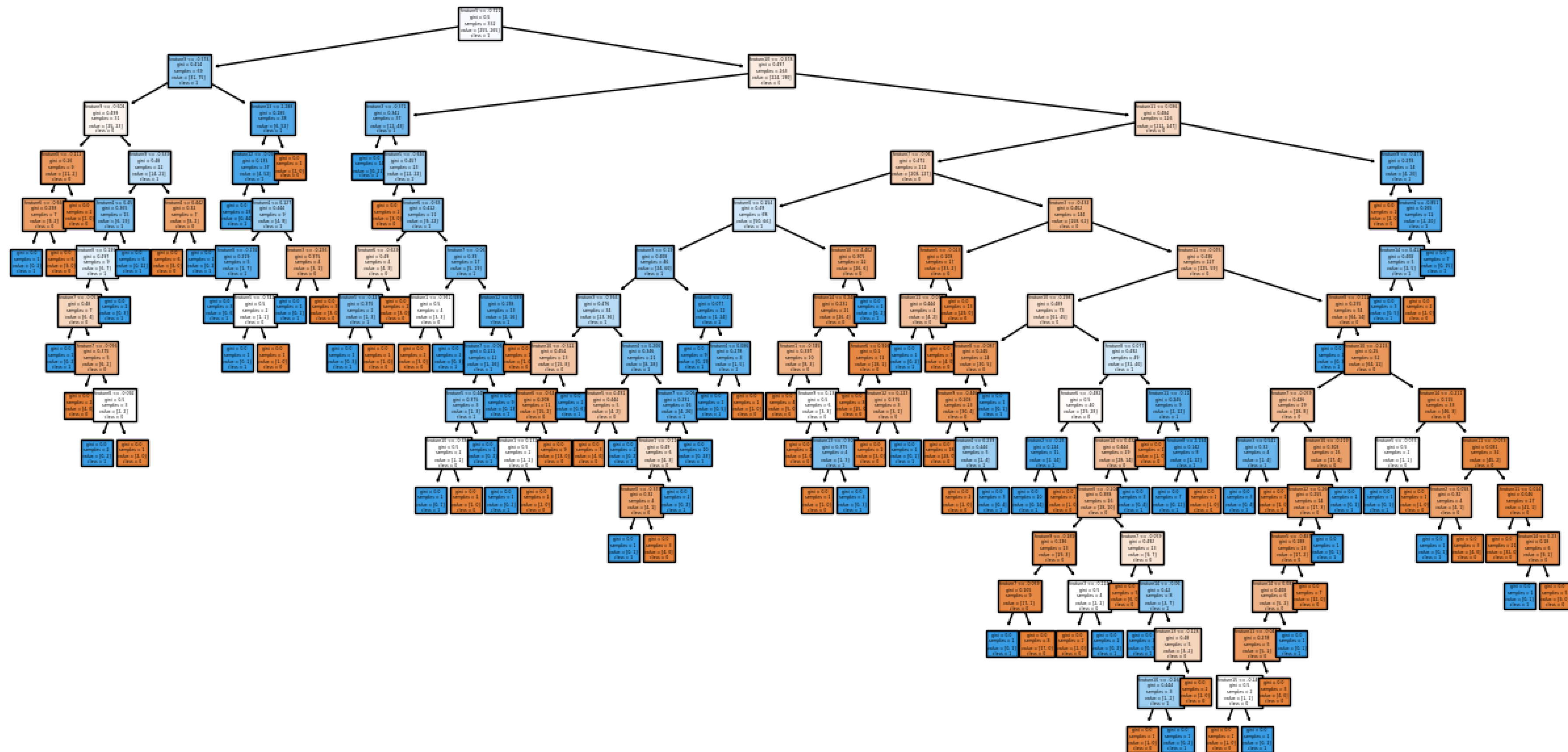
Importância das Características no Modelo Random Forest



Matriz de Confusão do Random Forest



ÁRVORE DE DECISÃO GERADA



CONCLUSÃO

COMPARAÇÃO ENTRE RANDOM FOREST E MLP NA CLASSIFICAÇÃO BINÁRIA

Razões para a Superioridade do Random Forest:

- Robustez e estabilidade;
- Manuseio de dados complexos;

Perspectivas Futuras:

- Otimização de hiperparâmetros;
- Exploração de uma base de dados maior;



REFERÊNCIAS

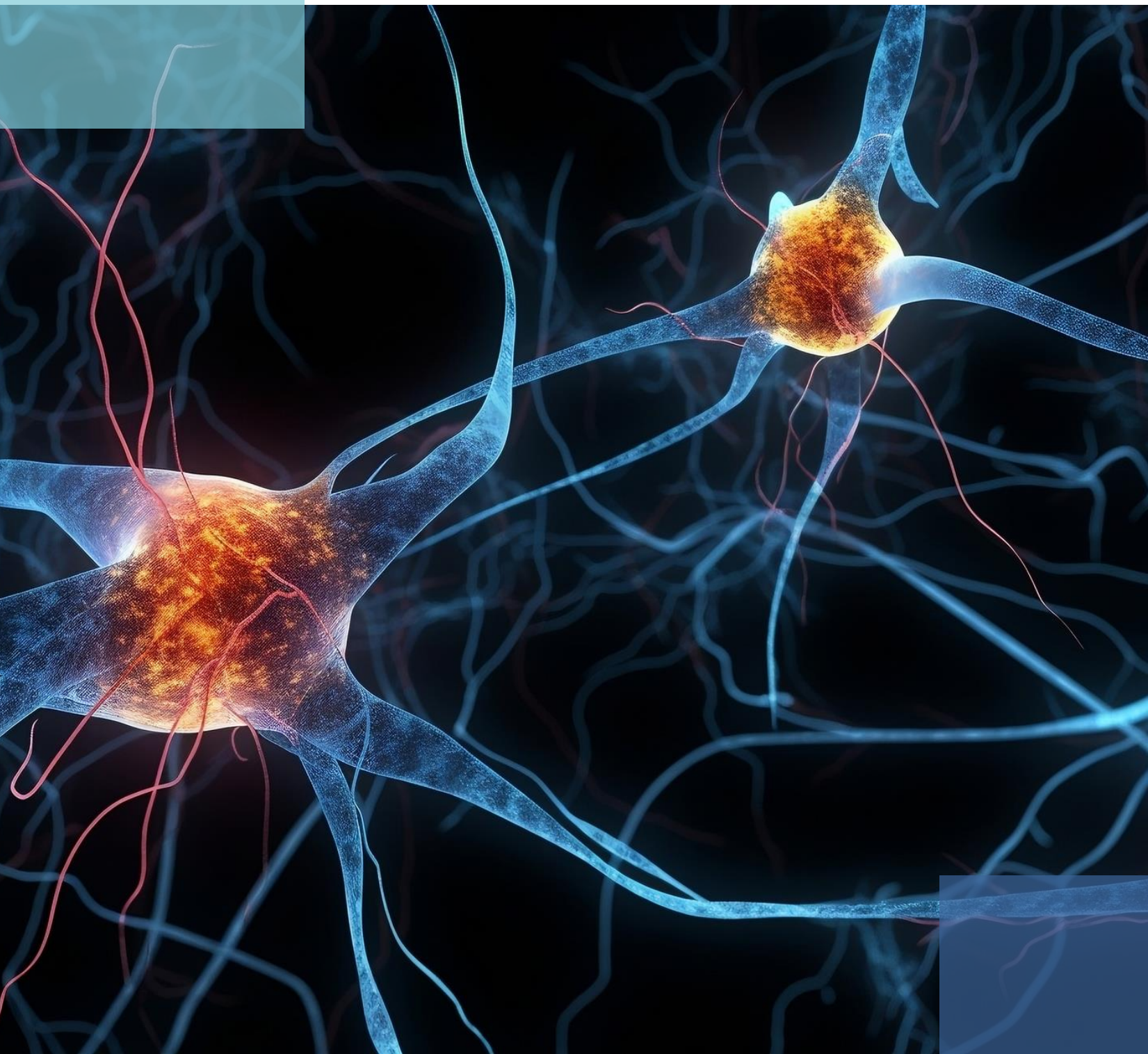
A. D. Kulkarni and B. Lowe, “Random forest algorithm for land cover classification,” International Journal on Recent and Innovation Trends in Computing and . . . , 2016.

F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.” Psychological review, vol. 65, no. 6, p. 386, 1958.

P.-N. Tan, M. Steinbach, and V. Kumar, “Data mining cluster analysis: basic concepts and algorithms,” Introduction to data mining, vol. 487, p. 533, 2013.

Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A survey of convolutional neural networks: analysis, applications, and prospects,” IEEE transactions on neural networks and learning systems, 2021.

R. C. Prati, G. Batista, M. C. Monard et al., “Curvas roc para avaliação de classificadores,” Revista IEEE América Latina, vol. 6, no. 2, pp. 215–222, 2008.



DÚVIDAS?

OBRIGADO
PELA ATENÇÃO!

