

# COMBINAÇÃO DE MODELOS E REDES NEURAIS ARTIFICIAIS

**PROF. DR. MURILO COELHO NALDI**

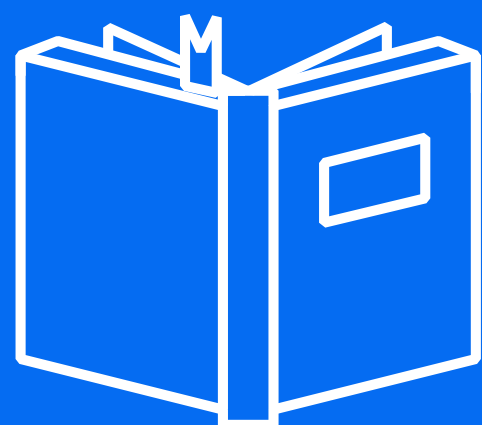
CARLOS EDUARDO FONTANELI  
RA 769949

IVAN DUARTE CALVO  
RA 790739



# Problemática

Aplicar as técnicas de combinação de modelos e de redes neurais artificiais em um dataset



## Conjunto de Dados

O objetivo é classificar, dado um conjunto de atributos relativos a um voo, se um passageiro ficará satisfeito ou não ao final da viagem.



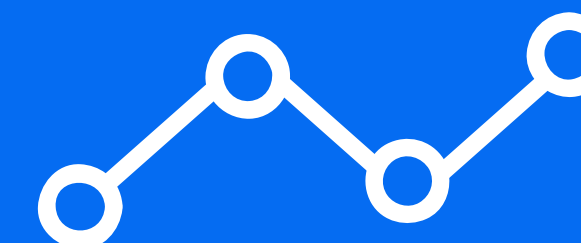
## Atributos

- Idade
- Sexo
- Classe do voo
- Distancia do voo
- entre outros



## Estudo dos Dados

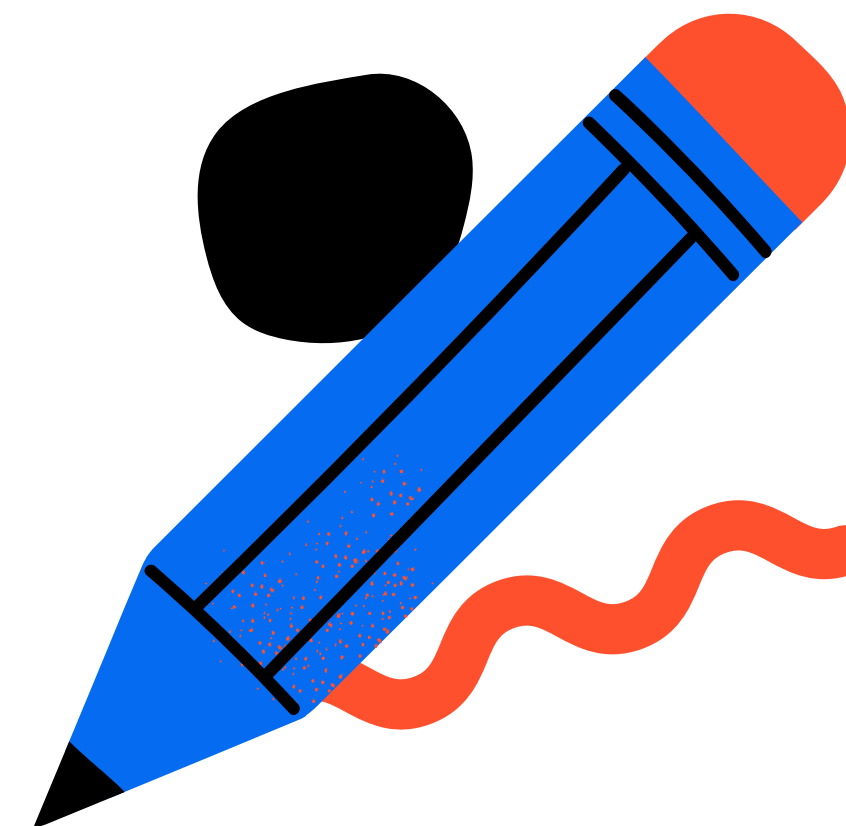
Realização da análise descritiva e exploratória dos dados.



## Treino, Teste e Avaliação de Modelos

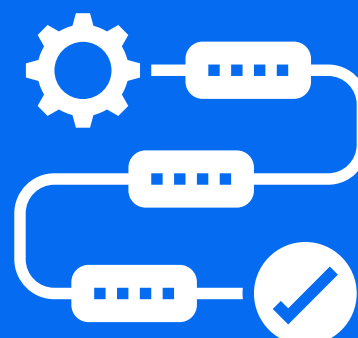
Implementação dos modelos de árvore de decisão, KNN, MLP, RNA, utilizando técnicas de combinação de modelos.

# Objetivos & Metodologia



## Objetivos

Gerar modelos preditivos e analisar o desempenho dos mesmos individualmente e depois em conjuntos.



## Metodologia

Análise exploratória dos dados.

Implementação sistematizada de modelos, com treino e teste sobre o conjunto de dados.

Avaliação dos resultados obtidos.



## Ferramentas

Linguagem de programação multi-paradigma orientada a objetos: Python.

Jupyter Notebooks para realização do relatório.

Canva para elaboração da apresentação.

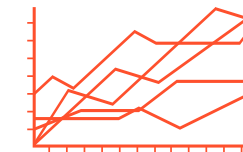
# Análise Exploratória



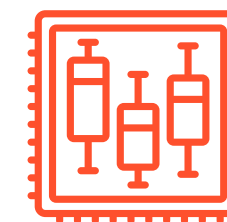
Obtenção de informações  
mais palpáveis e  
interpretativas



Busca de possíveis outliers  
e/ou desbalanceamentos



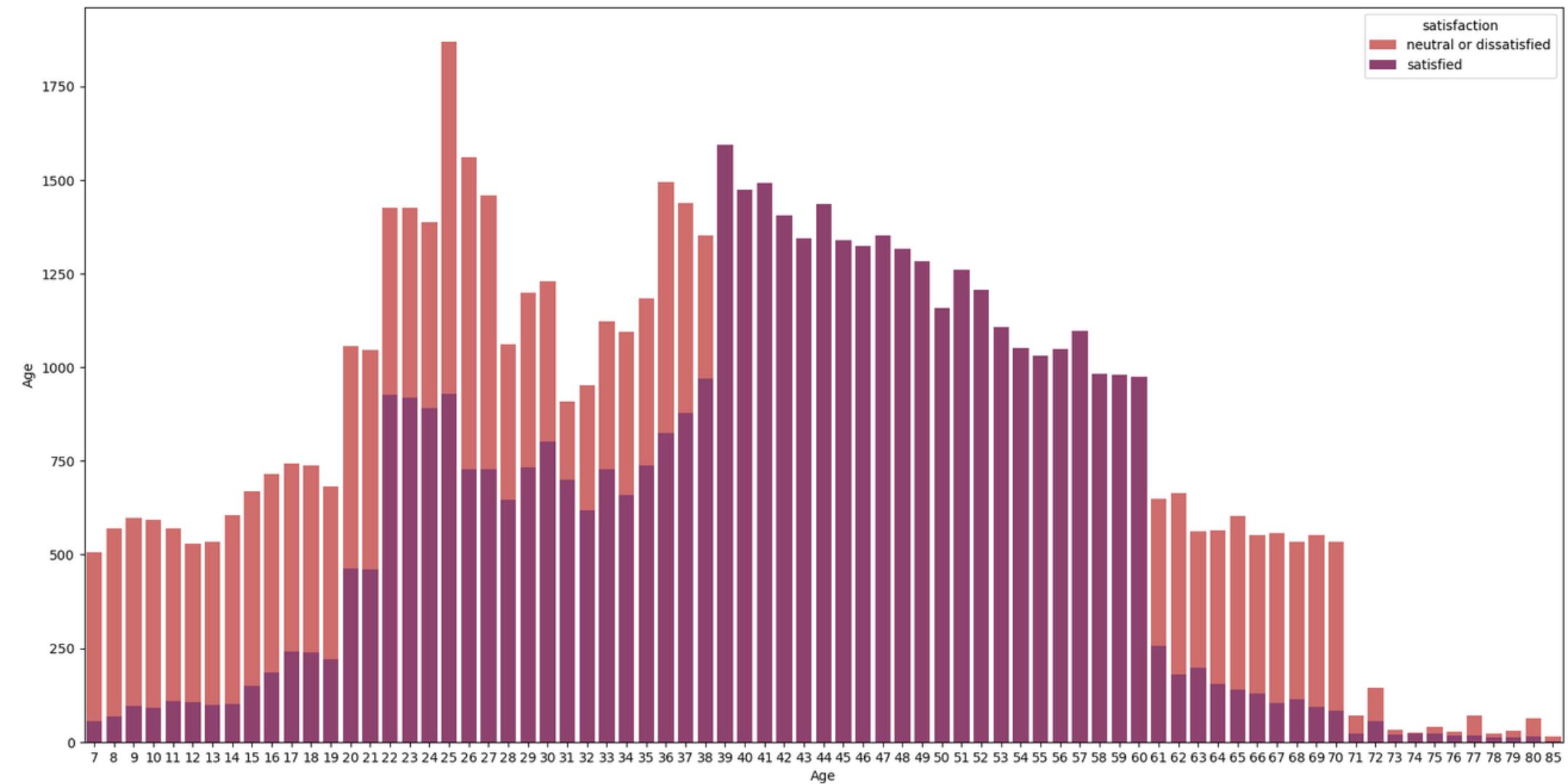
Representação visual dos  
dados através de gráficos



Boxplot das variáveis para  
representação de um  
conjunto de observações  
de uma variável  
quantitativa

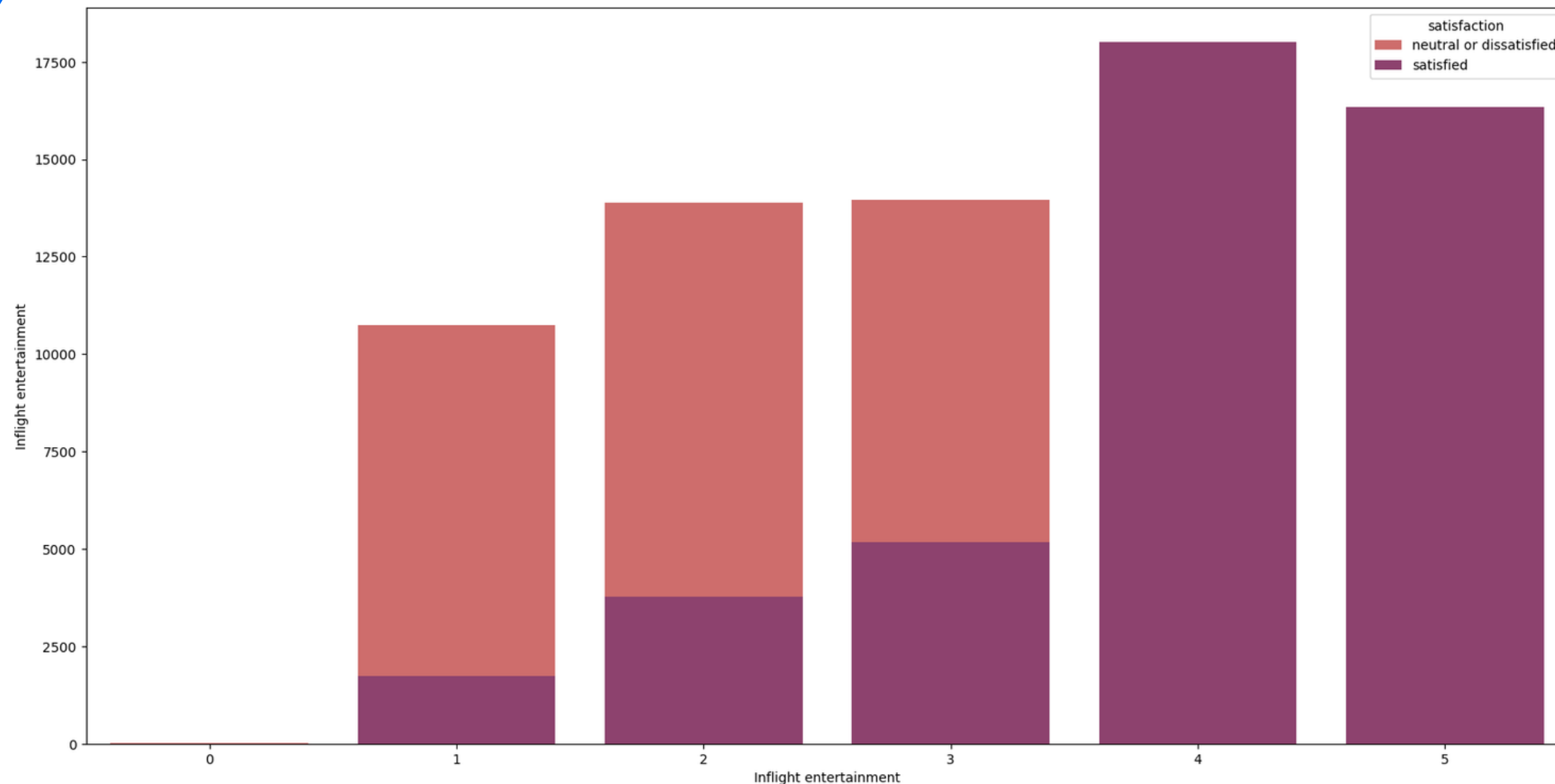
# Idade

- Clientes com faixa etária entre 40 e 59 anos tendem a ficar mais satisfeitos durante os voos;



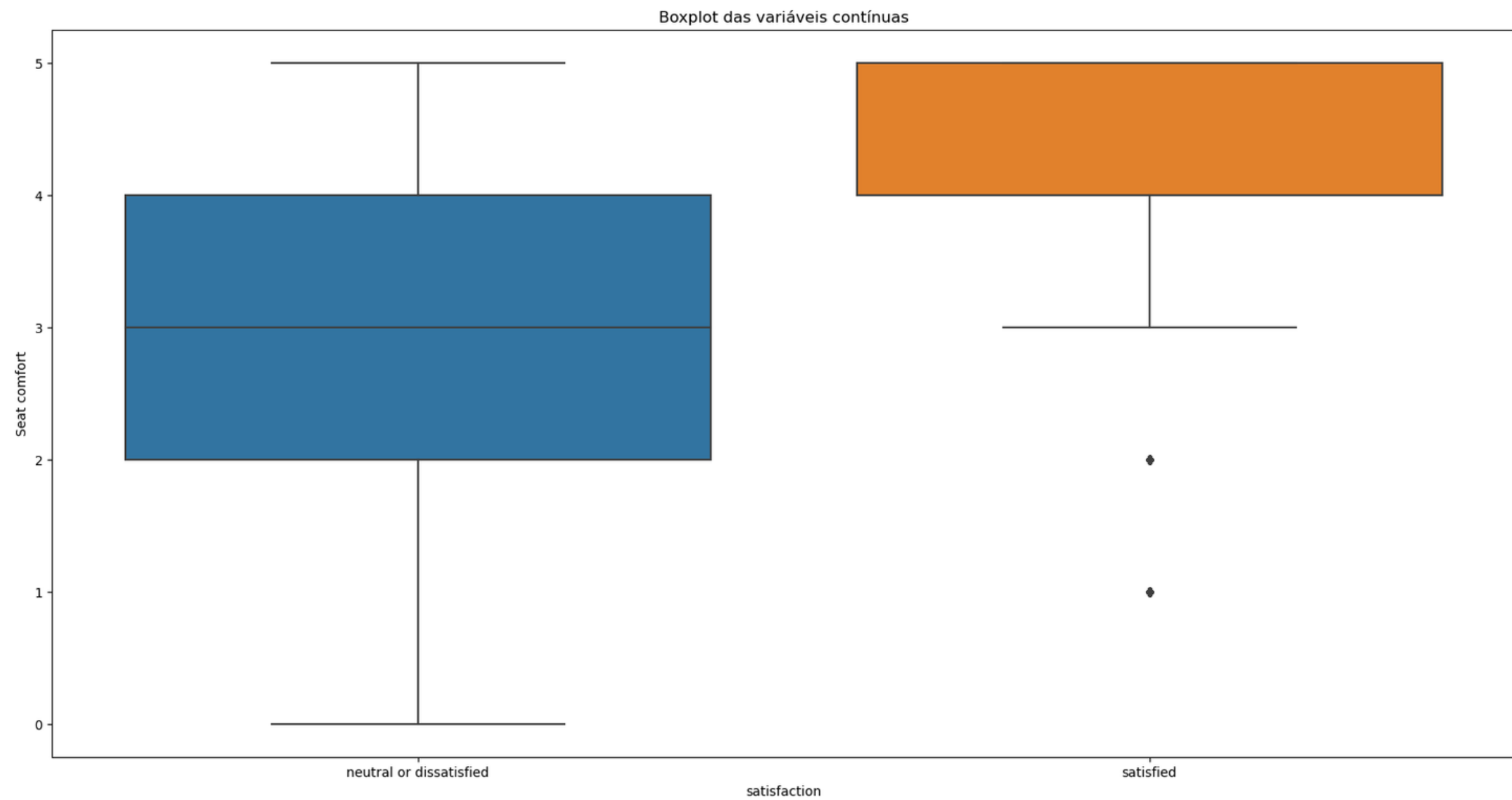
# Serviços durante o voo

- A qualidade dos serviços ofertados durante o voo está fortemente correlacionada com a satisfação do cliente;



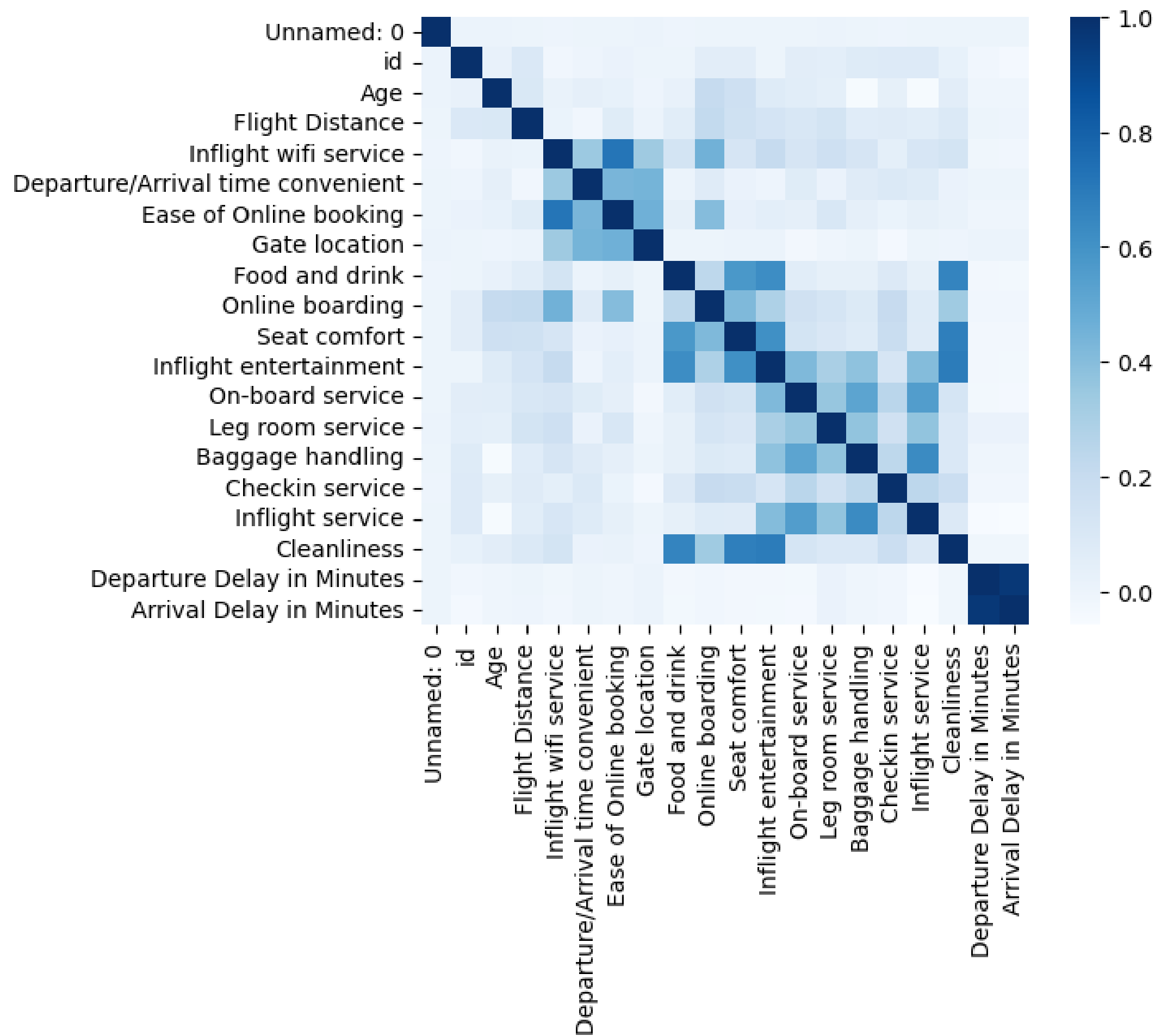
# Outliers

- Foram notados a presença de outliers relativo a clientes satisfeitos mesmo com serviços ofertados em pior qualidade;
- O oposto não foi observado;



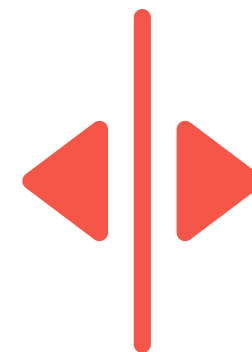
## 5

- 

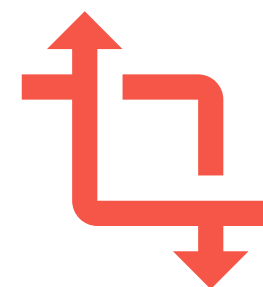




# Pré-Processamento



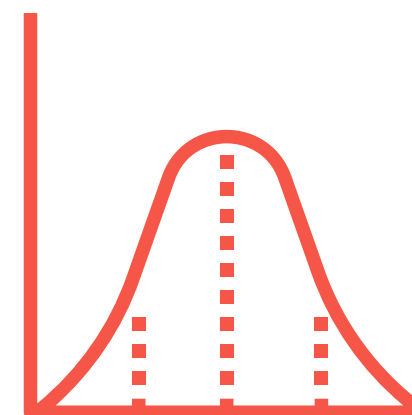
Separação entre atributos e alvo;



Discretização dos atributos nominais ( $[0, \text{número\_de\_classes} - 1]$ );

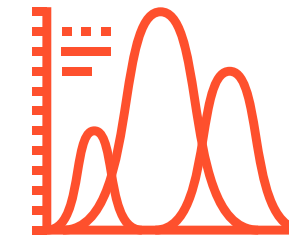
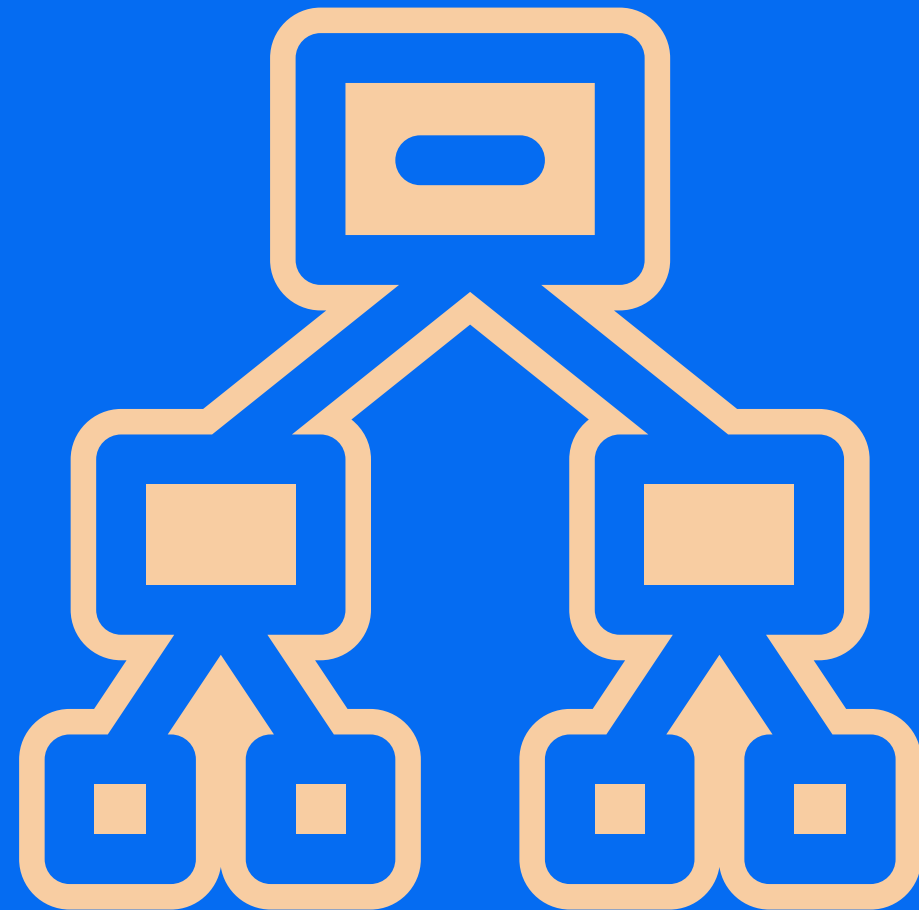


Remoção de outliers;

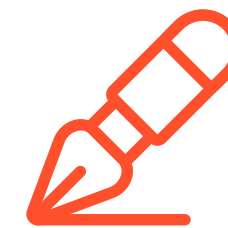


Normalização dos atributos;

# Árvore de Decisão



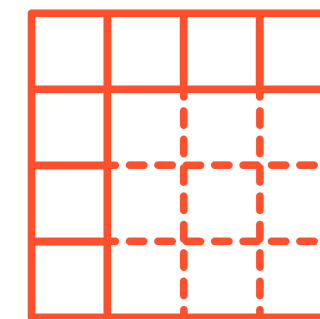
Divisão do conjunto de dados baseados em condições



Treinamento, teste e avaliação



Cálculo de métricas de avaliação



Análise visual dos resultados com matriz de confusão

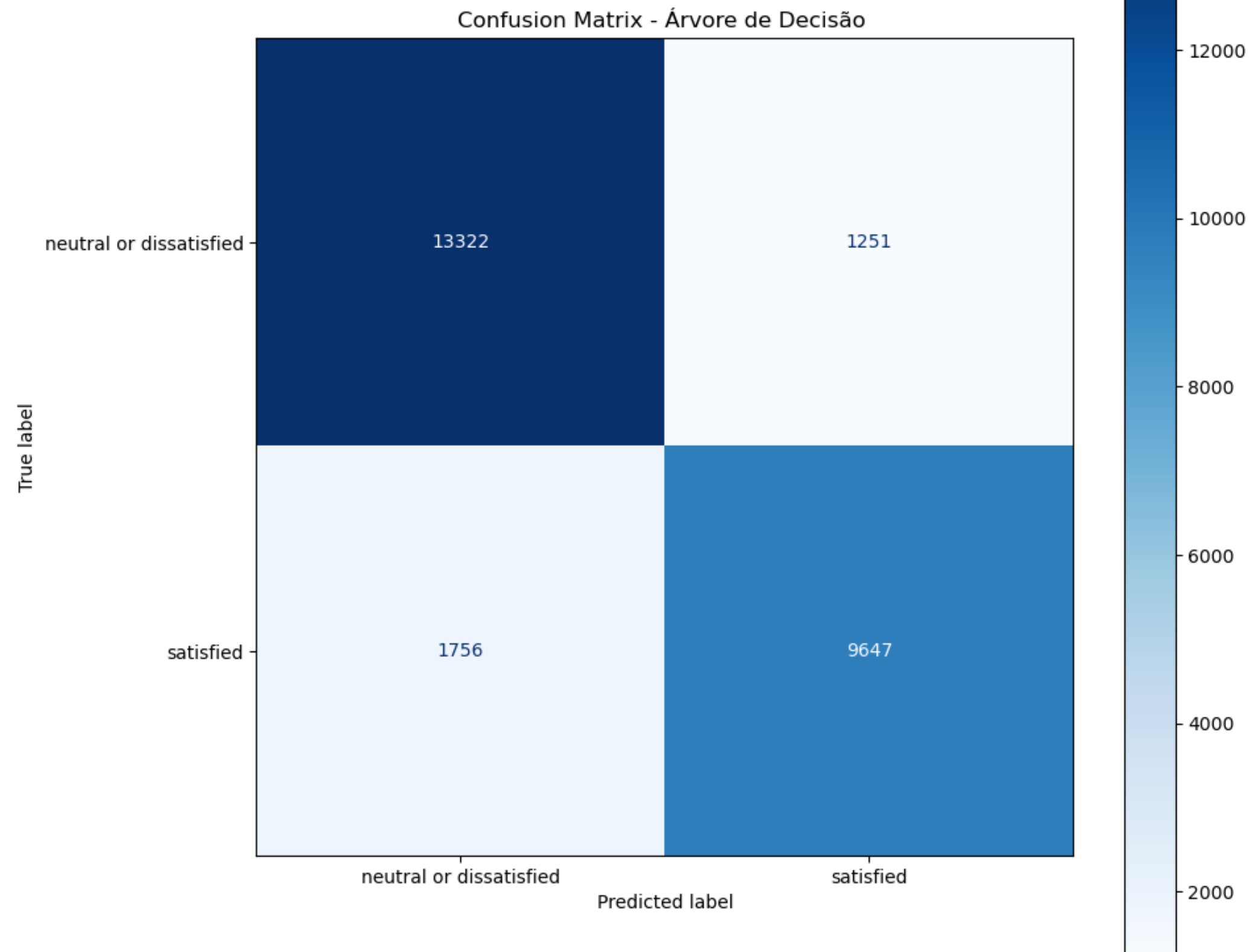
# Árvore de Decisão

- Acurácia balanceada de 88%
- Precisão de 88%

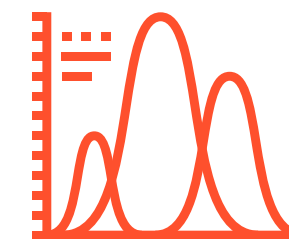
	precision	recall	f1-score	support
0	0.91	0.88	0.90	15078
1	0.85	0.89	0.87	10898
accuracy			0.88	25976
macro avg	0.88	0.88	0.88	25976
weighted avg	0.89	0.88	0.88	25976
0.8800808768206662				

# Árvore de Decisão

- Matriz de Confusão



# K-Vizinhos



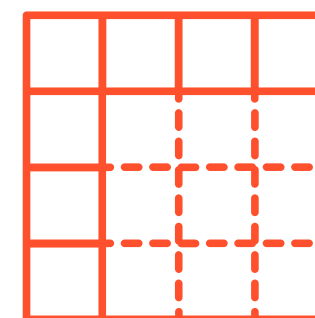
Implementa classificação através das distâncias



Treinamento, teste e avaliação



Cálculo de métricas de avaliação



Análise visual dos resultados com matriz de confusão

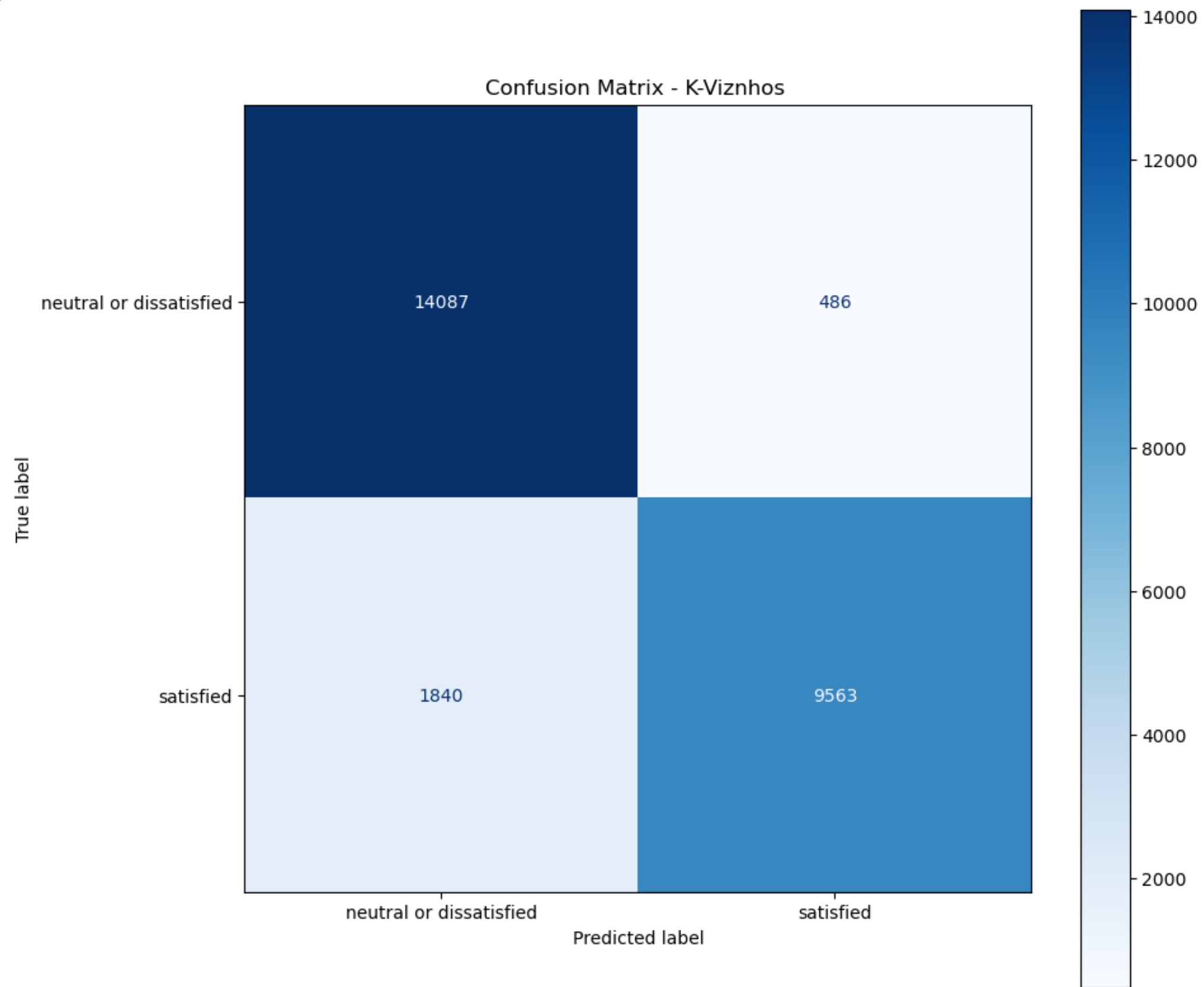
# K-Vizinhos

- Acurácia balanceada de 90%;
- k=5 (mas sem muitas alterações quando diferente)

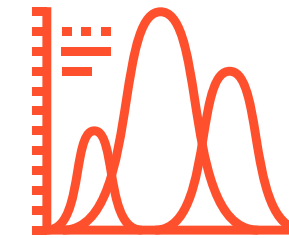
	precision	recall	f1-score	support
0	0.97	0.88	0.92	15927
1	0.84	0.95	0.89	10049
accuracy			0.91	25976
macro avg	0.90	0.92	0.91	25976
weighted avg	0.92	0.91	0.91	25976
0.9026448049912695				

# K-Vizinhos

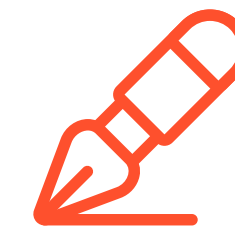
- Matriz de Confusão



# Perceptron Multicamadas



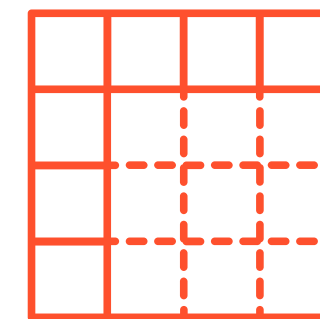
Rede de camadas de neurônios



Treinamento, teste e avaliação



Cálculo de métricas de avaliação



Análise visual dos resultados com matriz de confusão



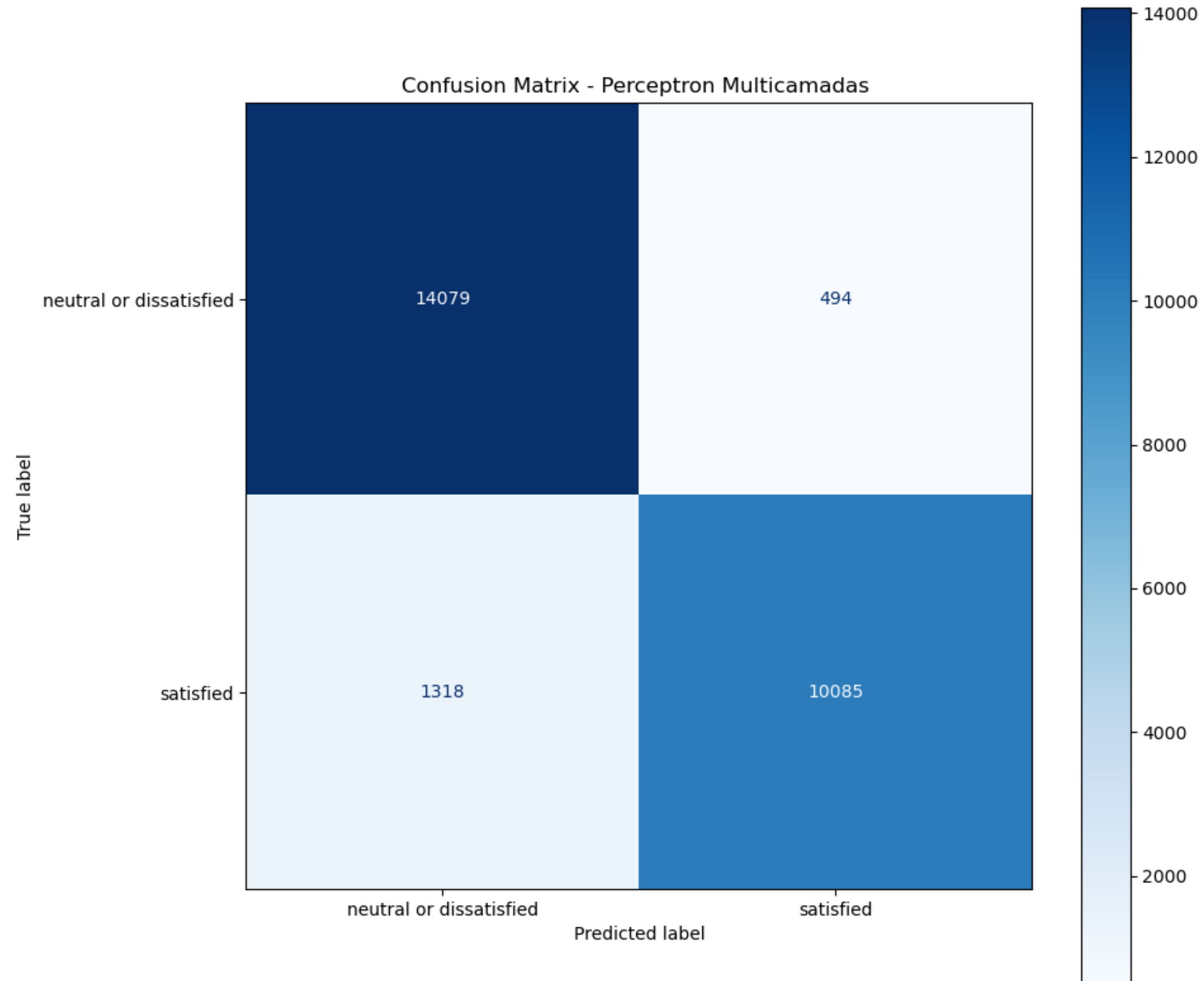
# Perceptron Multicamadas

- Acurácia balanceada de 92.5%;
- Precisão de 93%
- 100 hidden layers

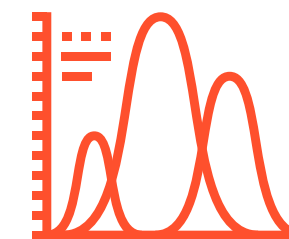
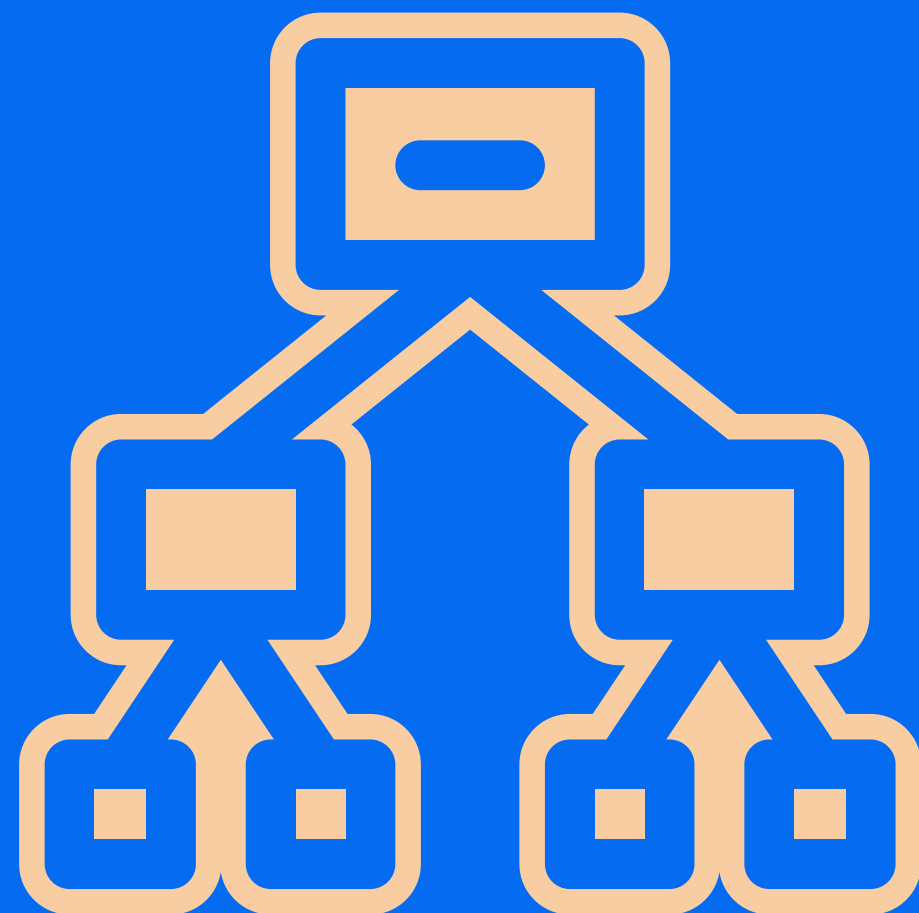
	precision	recall	f1-score	support
0	0.97	0.91	0.94	15397
1	0.88	0.95	0.92	10579
accuracy			0.93	25976
macro avg	0.93	0.93	0.93	25976
weighted avg	0.93	0.93	0.93	25976
0.9252590382846024				

# Perceptron Multicamadas

- Matriz de Confusão



# Rede Neural Artificial



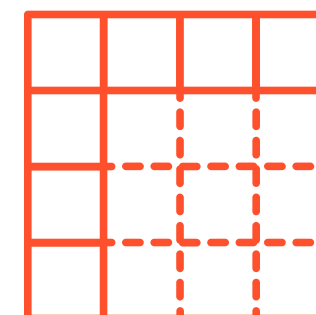
Implementação manual utilizando a biblioteca torch como mostrado em aula



21 camadas esconhidas, funções de ativação sigmoide e ReLU.



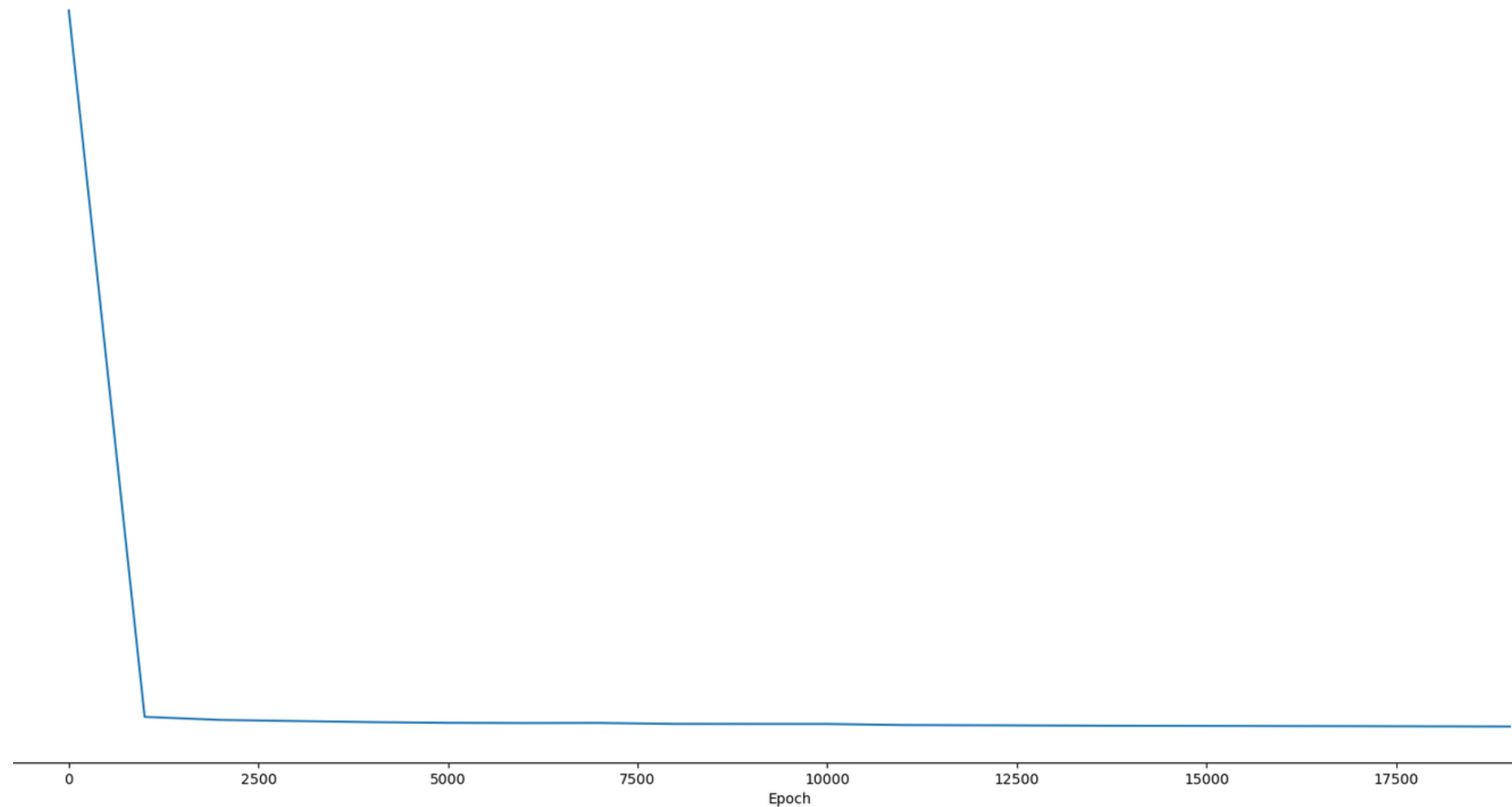
Treinamento em 20mil épocas, teste e avaliação,. Cálculo de métricas de avaliação



Análise visual dos resultados com matriz de confusão

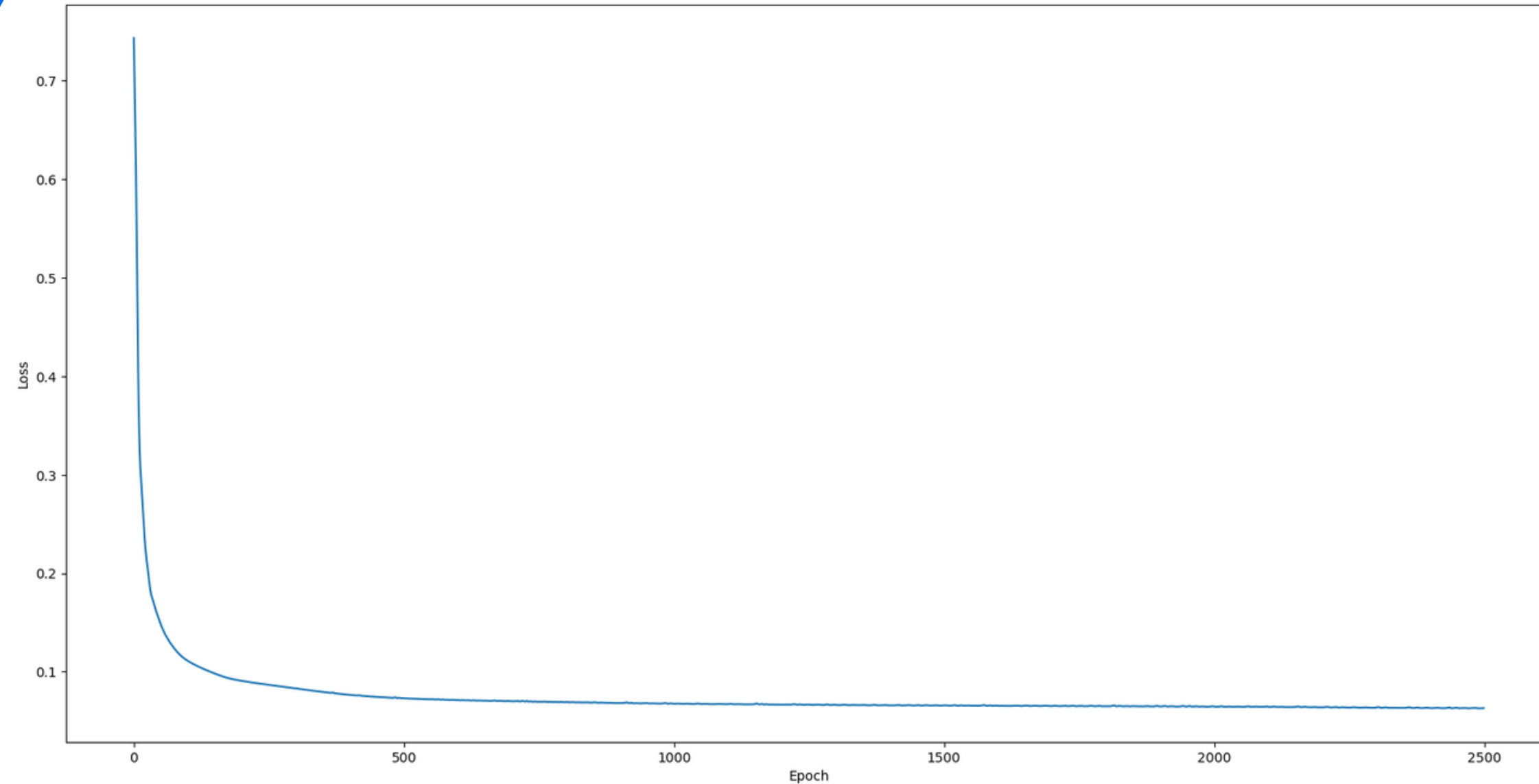
# Rede Neural Artificial

- A partir de 2500 épocas a queda se mostrou muito pequena;



# Rede Neural Artificial

- A partir de 2500 épocas a queda se mostrou muito pequena;



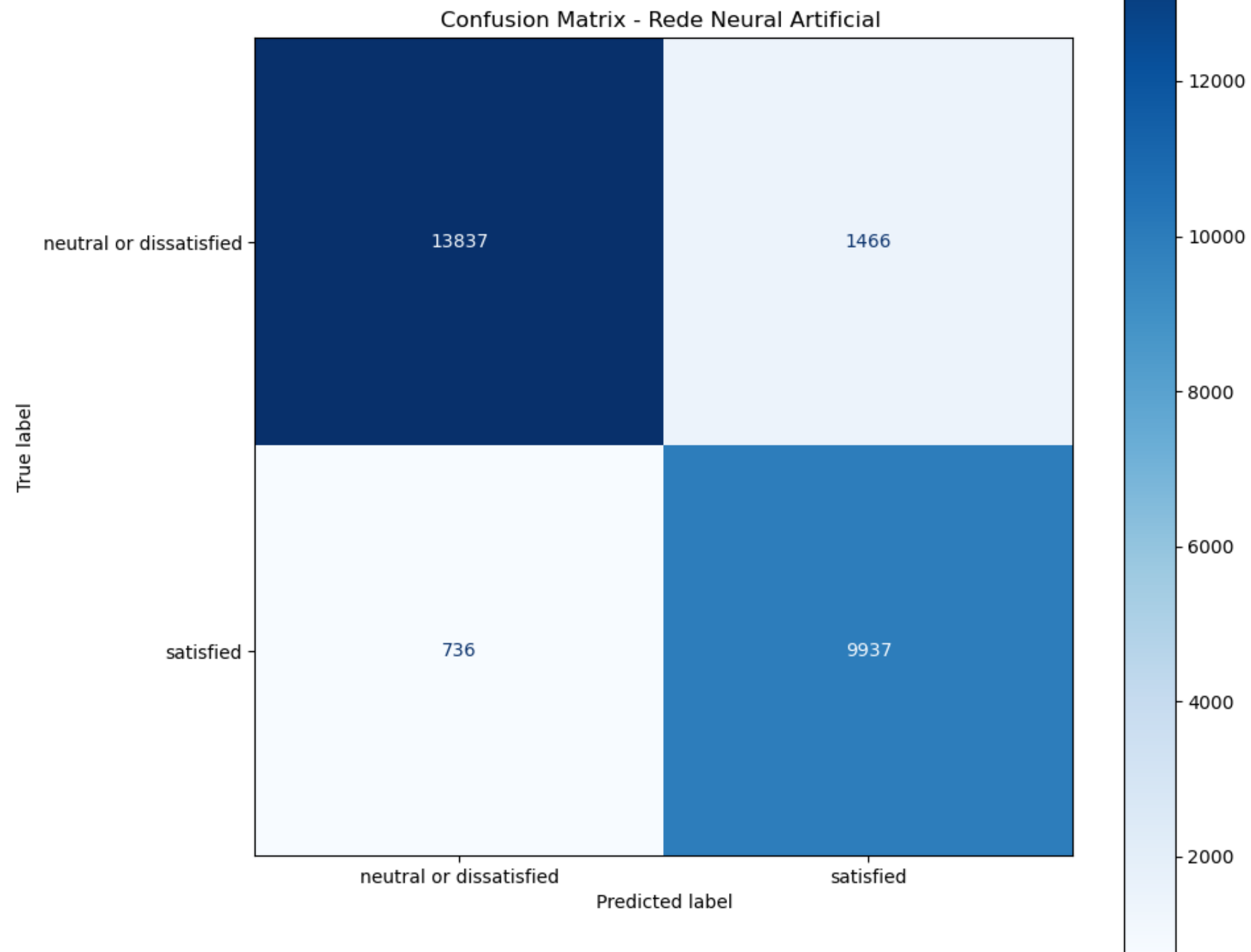
# Rede Neural Artificial

- Acurácia balanceada de 91.8%
- Precisão de 92%

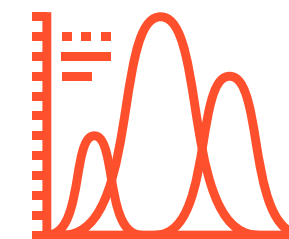
	precision	recall	f1-score	support
0	0.95	0.90	0.93	15303
1	0.87	0.93	0.90	10673
accuracy			0.92	25976
macro avg	0.91	0.92	0.91	25976
weighted avg	0.92	0.92	0.92	25976
0.9176213674689171				

# Rede Neural Artificial

- Matriz de Confusão



# Combinação



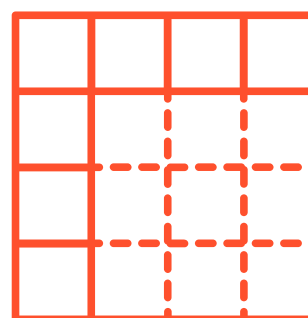
Junção de dois ou mais classificadores



Votação Majoritária



Votação por Scores



Análise visual dos resultados com matriz de confusão



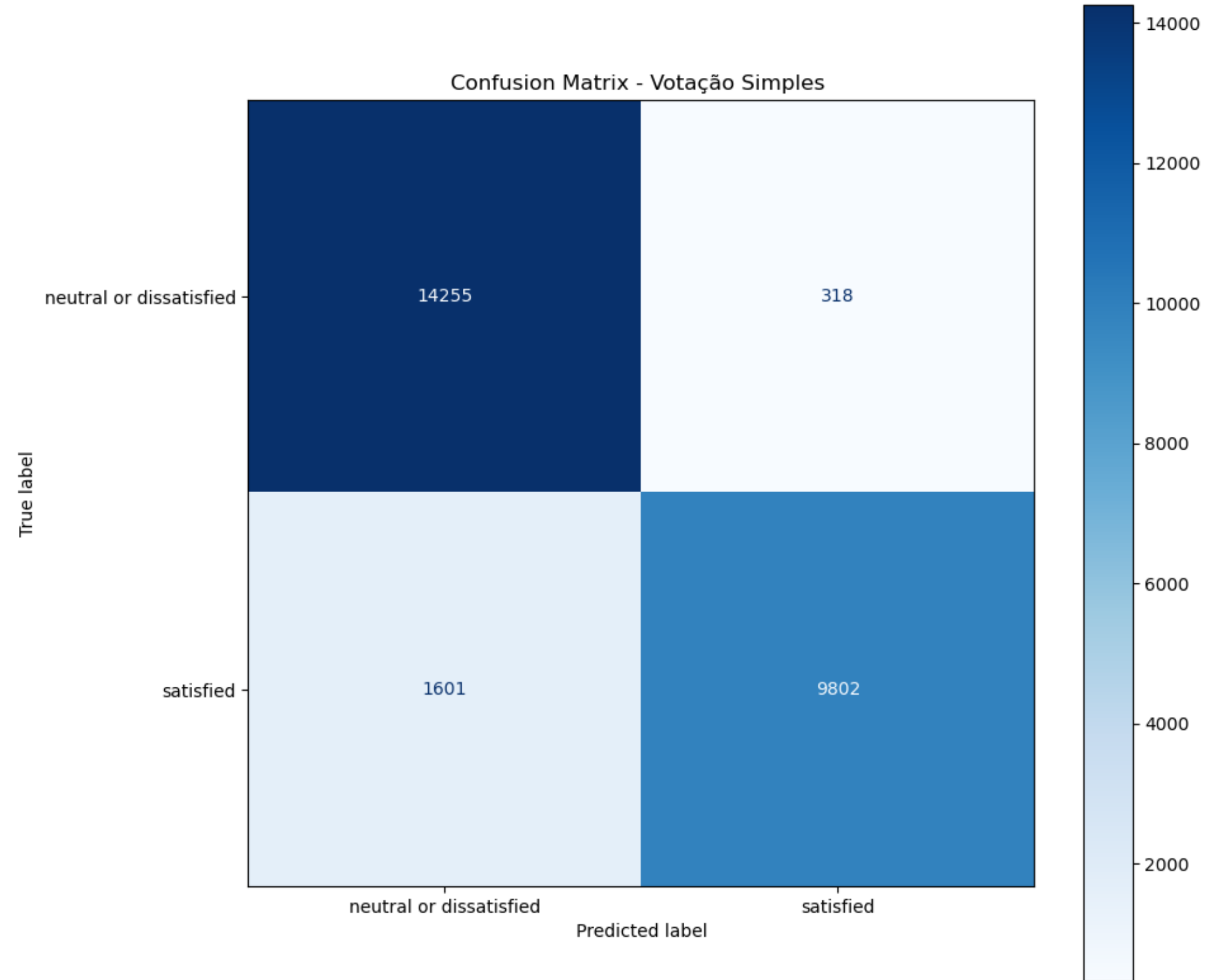
# Votação Majoritária

- Escolha com base na classe mais "votada"
- Acurácia balanceada de 91.9%

	precision	recall	f1-score	support
0	0.98	0.90	0.94	15856
1	0.86	0.97	0.91	10120
accuracy			0.93	25976
macro avg	0.92	0.93	0.92	25976
weighted avg	0.93	0.93	0.93	25976
0.9188885875816941				

# Votação Majoritária

- Escolha com base na classe mais "votada"
- Acurácia balanceada de 91.9%



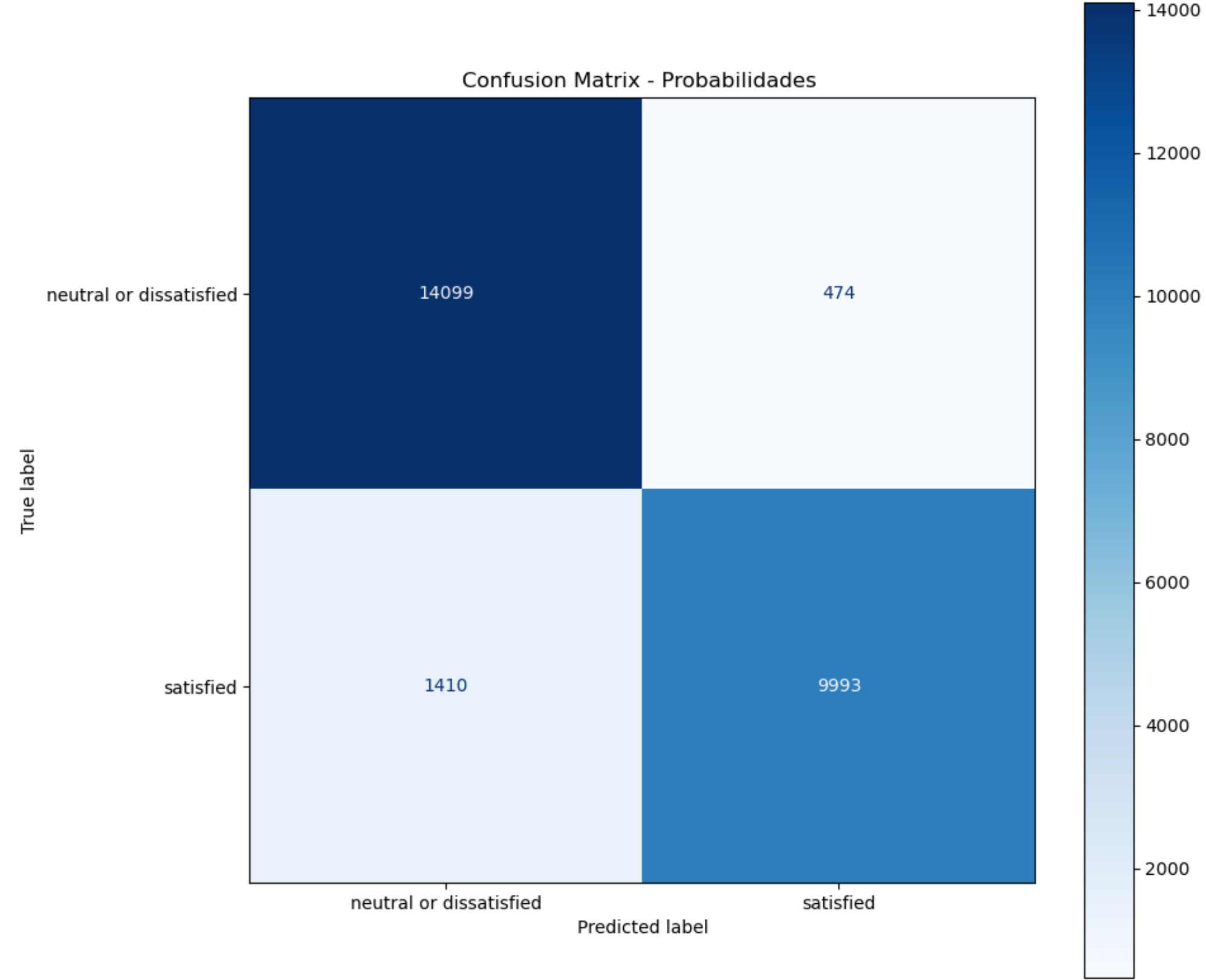
# Votação por probabilidade

- Escolha com base em scores
- Acurácia balanceada de 92%

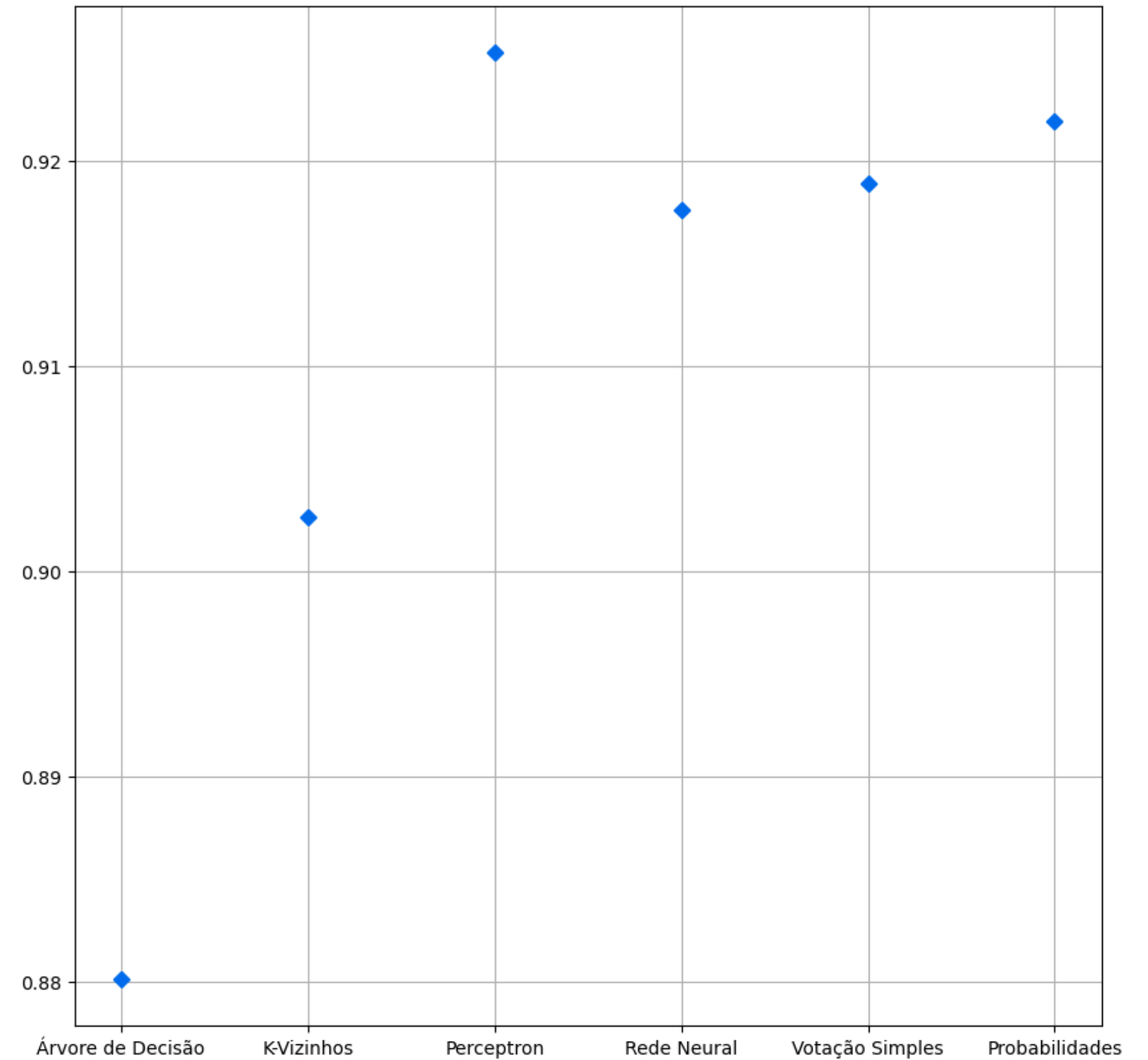
	precision	recall	f1-score	support
0	0.97	0.91	0.94	15509
1	0.88	0.95	0.91	10467
accuracy			0.93	25976
macro avg	0.92	0.93	0.93	25976
weighted avg	0.93	0.93	0.93	25976
0.9219112126589173				

# Votação por probabilidade

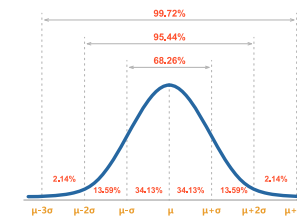
- Escolha com base em scores
- Acurácia balanceada de 92%



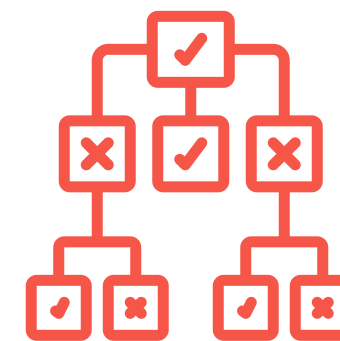
# Considerações Finais



# Considerações Finais



A Combinação propaga a predominância nos falsos negativos



Perceptron Multicamadas e Rede Neural atingem maiores eficácias que os outros classificadores



Combinação obtém bons resultados com uma possível variância maior

