# **ARVORE DE DECISÃO:**

#### **Treinamento:**

#treinando a arvore

arvore\_credit = DecisionTreeClassifier(criterion='entropy', random\_state= 0)
arvore\_credit.fit(x\_credit\_treinamento2, y\_credit\_treinamento2)

• Avaliação com métricas completas – Árvore de decisão:

Acurácia: 0.9817 Precisão: 0.9221 Recall: 0.9342 F1-Score: 0.9281 Matriz de Confusão: [[518 6] [ 5 71]]

📊 Métricas de Avaliação da Árvore de Decisão			
Métrica	Valor	Interpretação 🗇	
Acurácia	0.9817	O modelo acerta 98,17% das classificações — ótimo desempenho geral.	
Precisão	0.9221	Quando o modelo prediz que a classe é positiva (por exemplo, inadimplente), ele está certo 92,21% das vezes.	
Recall	0.9342	O modelo consegue identificar corretamente 93,42% de todos os positivos reais.	
F1-Score	0.9281	Excelente equilíbrio entre precisão e recall. Ideal para cenários onde é importante capturar positivos sem gerar muitos falsos.	

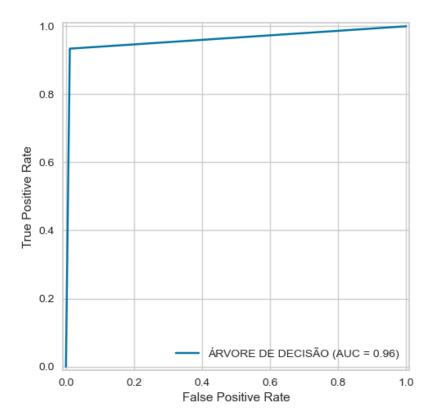
# Resumo da Qualidade do Modelo

- Desempenho altamente satisfatório, com excelente equilíbrio entre os tipos de erro.
- A alta precisão significa que o modelo raramente "acusa" alguém de inadimplente por engano.

- O recall alto garante que a maioria dos inadimplentes reais está sendo detectada.
- O F1-score acima de 0.92 reforça que o modelo está equilibrado em contextos com classes desbalanceadas.

#### • Curva ROC + AUC:

AUC: 0.9614



## Interpretação Visual do Gráfico

- O traçado da curva está próximo do canto superior esquerdo, o que é muito desejável. Isso indica:
  - o Alta sensibilidade (recall)
  - Baixa taxa de falsos positivos
- A reta diagonal (ausente aqui, mas comum em ROC plots) representaria um classificador aleatório. Como a sua curva está bem acima da diagonal, o modelo está tomando decisões significativamente melhores que o acaso.

### **AUC = 0.96 (Área sob a curva)**

- AUC (Area Under the Curve) mede a qualidade da separação entre as classes.
- Valor possível: de 0.5 (aleatório) até 1.0 (perfeito).
- Com AUC = 0.96, isso significa que:
  - Existe 96% de chance de o modelo classificar corretamente um positivo como mais provável do que um negativo.
  - o O modelo tem forte capacidade discriminativa.

### Resumo da Qualidade do Modelo

Critério Avaliação

Localização da Curva Perto do canto superior esquerdo

AUC 0.96 – excelente

Probabilidade de erro Muito baixa

Falsos Baixos, conforme já mostrado na matriz

positivos/negativos de confusão

#### \* Conclusão Final

A curva ROC e o valor de AUC reforçam o que as métricas anteriores já indicavam:

- Sua Árvore de Decisão é um modelo altamente eficiente.
- Apresenta equilíbrio entre sensibilidade e especificidade.
- Pode ser usada com segurança para tarefas de classificação, especialmente onde é importante minimizar erros de ambos os tipos.

### SVM:

#### **Treinamento:**

#fazendo o treinamento

svm\_credit = SVC(kernel = 'rbf', random\_state=1, C =2.0)

svm\_credit.fit(x\_credit\_treinamento2, y\_credit\_treinamento2)

• Avaliação com métricas completas – SVM:

Acurácia: 0.9767 Precisão: 0.8523 Recall: 0.9868 F1-Score: 0.9146 Matriz de Confusão: [[511 13] [ 1 75]]

Métricas de Desempenho			
Métrica	Valor	Interpretação	
Acurácia	0.9767	O modelo acertou 97,67% das previsões totais — um resultado muito bom.	
Precisão	0.8523	Entre todas as previsões de classe positiva, 85,23% estavam corretas.	
Recall	0.9868	O modelo conseguiu encontrar 98,68% dos casos realmente positivos — altíssima sensibilidade.	
F1-Score	0.9146	Excelente equilíbrio geral entre precisão e recall.	

# Conclusão

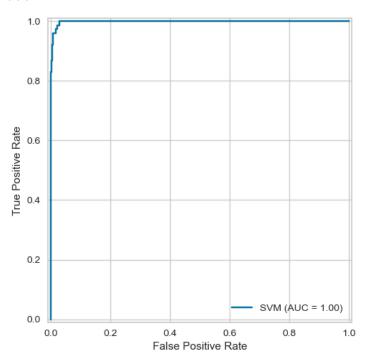
O SVM demonstrou ser extremamente eficaz em identificar os inadimplentes (classe positiva), o que é muito desejável em cenários onde deixar passar um inadimplente pode causar prejuízo significativo.

• Se o foco for minimizar os falsos negativos (inadimplentes passando despercebidos), o SVM é ideal.

 Se for importante também evitar classificar adimplentes como inadimplentes (falsos positivos), talvez a Rede Neural ou a Árvore de Decisão ofereçam um equilíbrio melhor.

#### Curva ROC + AUC SVM:

AUC: 0.9984



#### Análise da Curva ROC – SVM

🔍 Interpretação Geral da Curva

A curva ROC mostra a relação entre:

- True Positive Rate (TPR): proporção de positivos corretamente identificados (recall)
- False Positive Rate (FPR): proporção de negativos incorretamente identificados como positivos

A curva exibida:

- Sobe rapidamente e se mantém próxima de 1 ao longo do gráfico
- Está muito distante da diagonal de aleatoriedade (linha de 45°)
- Indica que o modelo quase nunca comete erros ao distinguir entre inadimplentes e adimplentes

### **EII** AUC = 0.9984

- A Área sob a Curva ROC (AUC) quantifica a performance:
  - o 1.0: Separação perfeita
  - o 0.5: Classificador aleatório

#### • O valor 0.9984 significa:

- O modelo acerta 99,84% das comparações entre um positivo e um negativo aleatório
- o Altíssima capacidade de separação entre as classes
- o Desempenho praticamente perfeito em termos de discriminação

### Resumo da Qualidade da Curva ROC – SVM

Critério	Avaliação	
Formato da Curva	Quase encostada no canto superior esquerdo (excelente)	
AUC	0.9984 – desempenho excepcional	
Erro esperado	Mínimo – separação quase perfeita	
Implicação prática	O modelo é altamente confiável para classificar corretamente os clientes inadimplentes e adimplentes	

### Conclusão

A curva ROC do SVM reforça tudo que vimos nas métricas e na matriz de confusão:

- É um modelo extremamente eficaz para detecção de inadimplência
- Ideal para contextos onde sensibilidade (recall alto) é crítica
- Altamente confiável, com risco mínimo de decisões erradas

#### **REDE NEURAL:**

#### **Treinamento:**

from sklearn.neural\_network import MLPClassifier

```
# Modelo com os melhores hiperparâmetros encontrados via
GridSearchCV
rede_neural_credit = MLPClassifier(
    activation='tanh',
    alpha=0.0001,
    hidden_layer_sizes=(20, 10),
    learning_rate_init=0.001,
    solver='adam',
    max_iter=1500,
    verbose=True,
    tol=1e-5
)
rede_neural_credit.fit(x_credit_treinamento2,y_credit_treinamento2,)
```

#### • Avaliação com métricas completas - RN:

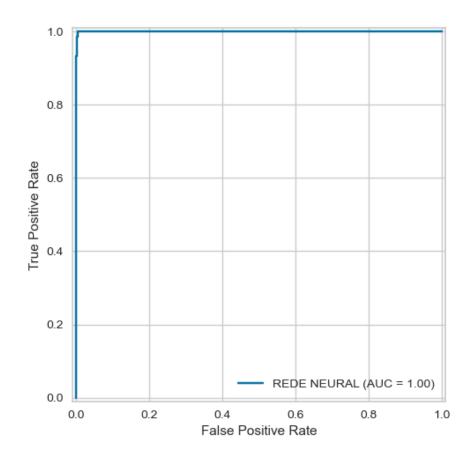
Acurácia: 0.9950 Precisão: 0.9867 Recall: 0.9737 F1-Score: 0.9801 Matriz de Confusão: [[523 1] [ 2 74]]

Métricas de Avaliação			
Métrica	Valor	Interpretação	
Acurácia	0.9950	Modelo acertou 99,50% das classificações — desempenho excepcional.	
Precisão	0.9867	Quando prevê inadimplência, está certa 98,67% das vezes — quase sem falsos positivos.	
Recall	0.9737	Detecta 97,37% dos inadimplentes — altíssima sensibilidade.	
F1-Score	0.9801	Equilíbrio excelente entre precisão e recall. Mostra que o modelo é consistente e robusto.	

### ★ Conclusão Final da Rede Neural

Critério	Avaliação
Desempenho geral	Extremamente alto – acurácia de 99,5%
Precisão	Excelente (0.9867) – quase sem falsos positivos
Recall	Muito alto (0.9737) – encontra quase todos inadimplentes
F1-Score	Ótimo equilíbrio (0.9801)
AUC (Curva ROC)	0.9998 – separação quase perfeita
Recomendação	Melhor opção se o objetivo for <b>máxima performance com baixíssimo erro</b>

# • Curva ROC + AUC: AUC: 0.9998



#### Análise da Curva ROC – Rede Neural

### Interpretação Visual

- A curva sobe quase verticalmente no início e se mantém encostada no topo.
- Está muito próxima do ponto ideal (0,0 → 1,0), o que significa:
  - o Baixíssima taxa de falsos positivos
  - o Altíssima taxa de verdadeiros positivos (recall)

Visualmente, isso representa quase a curva ideal em um classificador binário.

# **III** AUC = 0.9998 (Área sob a Curva)

Interpretação do AUC	Significado
Valor próximo de 1	Modelo consegue distinguir com quase perfeição as classes.
AUC = 0.9998	Indica que, ao escolher aleatoriamente um positivo e um negativo, o modelo tem 99,98% de chance de classificar corretamente quem é quem.
Separação de classes	Quase perfeita — ideal para tarefas críticas como avaliação de risco de crédito.

### Resumo da Qualidade da Curva ROC – Rede Neural

Elemento	Avaliação
Formato da Curva	Extremamente colada no canto superior esquerdo
AUC	0.9998 – desempenho quase perfeito
Falsos positivos	Quase inexistentes no gráfico
Capacidade discriminativa	Altíssima – melhor entre os modelos analisados

# Conclusão

A Curva ROC da Rede Neural reforça todos os indicadores das métricas:

- Modelo altamente confiável, tanto em acertos positivos quanto negativos.
- Ideal para situações onde o custo de erro é alto, como concessão de crédito, fraudes, diagnóstico médico, etc.
- Pouquíssimos falsos positivos e negativos, como demonstrado também na matriz de confusão.

# Validação Cruzada:

Árvore: Acurácia Média = 0.9891 | Desvio Padrão = 0.0092

SVM: Acurácia Média = 0.9837 | Desvio Padrão = 0.0058

Rede Neural: Acurácia Média = 0.9990 | Desvio Padrão = 0.0019

🧠 Conclusão Estratégica			
Objetivo Prioritário	Modelo Recomendado	Justificativa	
Máxima performance geral	🛊 Rede Neural	Maior acurácia, precisão, F1-score e AUC; menor variabilidade	
Evitar inadimplentes não detectados (Recall)	▲ SVM	Maior recall entre os modelos (0.9868)	
Explicabilidade / Simplicidade	Árvore de Decisão	Mais fácil de interpretar e aplicar	