

Novas métricas para classificação: ROC e AUC



Relembrando das métricas da última aula...

No caso das operações financeiras, chegamos em

CLASSE VERDADEIRA			
		1	0
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767 → 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119 → 55127
		↓ 75	↓ 56886 → 56961

$$\text{Taxa de acerto} = \frac{VP+VN}{VP+VN+FP+FN} = \frac{67+55119}{67+55119+1767+8} = 96,7\%$$

$$\text{precisão} = \frac{VP}{VP+FP} = \frac{67}{67+1767} = 3,6\%$$

Métrica que penaliza a ocorrência de **FP**

$$\text{recall} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

Métrica que penaliza a ocorrência de **FN**

$$\text{F1 score} = \frac{2(\text{precisão})(\text{recall})}{(\text{precisão})+(\text{recall})} = 6,9\%$$

Métrica que engloba as duas anteriores

Algumas novas considerações sobre a matriz de confusão

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

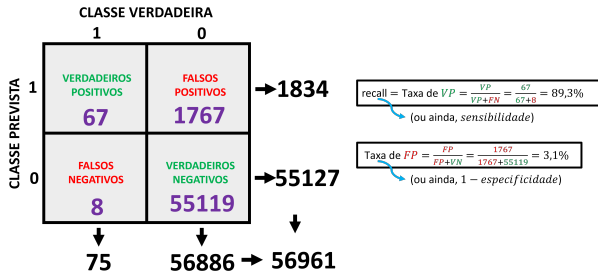
$$\text{recall} = \text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, *sensibilidade*)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, $1 - \text{especificidade}$)

- ❶ O *recall* (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.



- 1 O *recall* (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- 2 Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ → 56961

$$\text{recall} = \text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, sensibilidade)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, 1 - especificidade)

- 1 O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- 2 Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- 3 Quanto maior a **taxa de verdadeiros positivos**, melhor.

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

$$\text{recall} = \text{Taxa de } VP = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, *sensibilidade*)

$$\text{Taxa de } FP = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, $1 - \text{especificidade}$)

- ❶ O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- ❷ Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- ❸ Quanto maior a **taxa de verdadeiros positivos**, melhor.
- ❹ Analogamente, também podemos obter a chamada **taxa de falsos positivos**, sendo ela calculada pela expressão mostrada acima.

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

$$\text{recall} = \text{Taxa de } VP = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, sensibilidade)

$$\text{Taxa de } FP = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, 1 - especificidade)

- ❶ O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- ❷ Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- ❸ Quanto maior a **taxa de verdadeiros positivos**, melhor.
- ❹ Analogamente, também podemos obter a chamada **taxa de falsos positivos**, sendo ela calculada pela expressão mostrada acima.
- ❺ Para este exemplo específico, ela representa a quantidade de operações incorretamente classificadas como fraudulentas em relação ao total de operações não fraudulentas.

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

$$\text{recall} = \text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, sensibilidade)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, 1 - especificidade)

- 1 O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- 2 Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- 3 Quanto maior a **taxa de verdadeiros positivos**, melhor.
- 4 Analogamente, também podemos obter a chamada **taxa de falsos positivos**, sendo ela calculada pela expressão mostrada acima.
- 5 Para este exemplo específico, ela representa a quantidade de operações incorretamente classificadas como fraudulentas em relação ao total de operações não fraudulentas.
- 6 Quanto maior a **taxa de falsos positivos**, pior.

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

$$\text{recall} = \text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, sensibilidade)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, 1 - especificidade)

- ❶ O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- ❷ Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- ❸ Quanto maior a **taxa de verdadeiros positivos**, melhor.
- ❹ Analogamente, também podemos obter a chamada **taxa de falsos positivos**, sendo ela calculada pela expressão mostrada acima.
- ❺ Para este exemplo específico, ela representa a quantidade de operações incorretamente classificadas como fraudulentas em relação ao total de operações não fraudulentas.
- ❻ Quanto maior a **taxa de falsos positivos**, pior.
- ❼ A **taxa de falsos positivos** é também chamada de "complemento da especificidade" (1 - especificidade)

		CLASSE VERDADEIRA		
		1	0	
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 67	FALSOS POSITIVOS 1767	→ 1834
	0	FALSOS NEGATIVOS 8	VERDADEIROS NEGATIVOS 55119	→ 55127
		↓ 75	↓ 56886	↓ 56961

$$\text{recall} = \text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{67}{67+8} = 89,3\%$$

(ou ainda, *sensibilidade*)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1767}{1767+55119} = 3,1\%$$

(ou ainda, *1 - especificidade*)

- ❶ O **recall** (revocação) é também chamado de **taxa de verdadeiros positivos**, ou ainda, de **sensibilidade**.
- ❷ Para este exemplo específico, ela representa a quantidade de transações corretamente classificadas como fraudulentas em relação à quantidade total de operações fraudulentas.
- ❸ Quanto maior a **taxa de verdadeiros positivos**, melhor.
- ❹ Analogamente, também podemos obter a chamada **taxa de falsos positivos**, sendo ela calculada pela expressão mostrada acima.
- ❺ Para este exemplo específico, ela representa a quantidade de operações incorretamente classificadas como fraudulentas em relação ao total de operações não fraudulentas.
- ❻ Quanto maior a **taxa de falsos positivos**, pior.
- ❼ A **taxa de falsos positivos** é também chamada de "complemento da especificidade" ($1 - \text{especificidade}$)
- ❽ Para esse caso acima, seria possível reduzir a **taxa de FP**, mantendo a **taxa de VP**? Nós veremos isso agora, estudando dois conceitos correlacionados entre si: **ROC** e **AUC**.

Métricas ROC e AUC

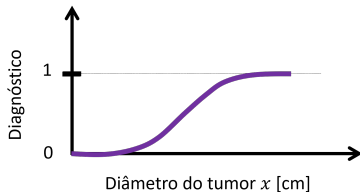
Métricas ROC e AUC

Começaremos com um exemplo mais simples, levando em conta um modelo de Regressão Logística.

- Considere que você tinha o **conjunto de dados de treinamento** ilustrado abaixo.
- Com esses dados de treinamento, um modelo de Regressão Logística foi treinado.



Como pode ser visto, o modelo resultante é conforme abaixo, e pode ser usado para classificar **novas amostras**



continua no próximo slide...

Suponha que você submeteu o modelo resultante a um **novo conjunto de dados** (dados de validação), utilizando um valor de limiar de 0.5 para realização da previsão.



Métricas ROC e AUC

Suponha que você submeteu o modelo resultante a um **novo conjunto de dados** (dados de validação), utilizando um valor de limiar de 0.5 para realização da previsão.



Analisando a figura, note que é possível construir a seguinte matriz de confusão resultante:

		CLASSE VERDADEIRA	
		1	0
CLASSE PREVISTA	1	VERDADEIROS POSITIVOS 3	FALSOS POSITIVOS 1
	0	FALSOS NEGATIVOS 1	VERDADEIROS NEGATIVOS 3

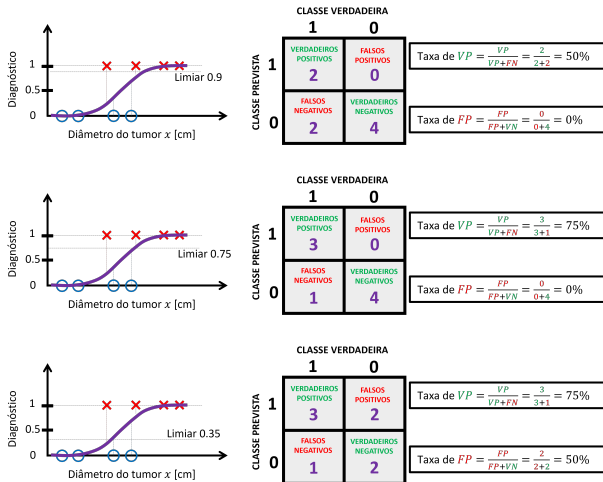
$$\text{Taxa de VP} = \frac{VP}{VP+FN} = \frac{3}{3+1} = 75\%$$

(= recall = sensibilidade)

$$\text{Taxa de FP} = \frac{FP}{FP+VN} = \frac{1}{1+3} = 25\%$$

(ou ainda, 1 - especificidade)

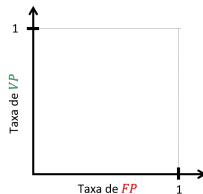
Para o mesmo conjunto de dados, note que, ao variar o valor de limiar, diferentes matrizes de confusão são obtidas.



- Observe que, à medida que o valor de limiar diminui, a taxa de verdadeiros positivos e a taxa de falsos positivos tendem a aumentar.

- Usando vários valores de limiar, podemos construir uma tabela com 3 colunas: valor de limiar, taxa de VP, taxa de FP
- A partir desta tabela, podemos construir um gráfico que relaciona a **taxa de verdadeiros positivos** com a **taxa de falsos positivos** para diferentes **valores de limiar**

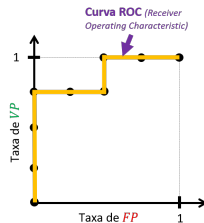
Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



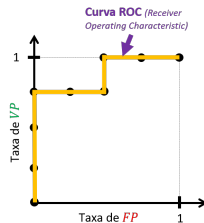
Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%

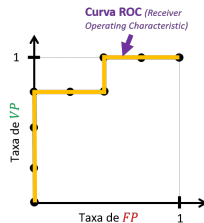


Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



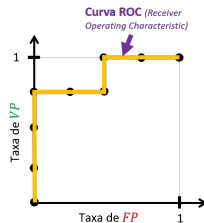
- A curva produzida é chamada de Curva ROC

Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



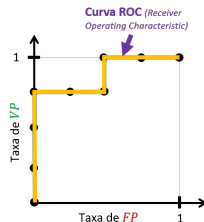
- A curva produzida é chamada de Curva ROC
- Pois bem, olhando para essa figura, qual é o melhor valor de limiar possível?

Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



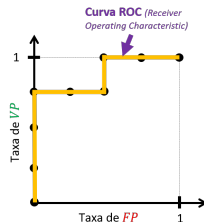
- A curva produzida é chamada de Curva ROC
- Pois bem, olhando para essa figura, qual é o melhor valor de limiar possível?
- Note que usar 0.2 é melhor do que usar 0 ou 0.05

Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



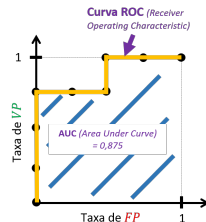
- A curva produzida é chamada de Curva ROC
- Pois bem, olhando para essa figura, qual é o melhor valor de limiar possível?
- Note que usar 0.2 é melhor do que usar 0 ou 0.05
- Note que usar 0.75 é melhor do que usar 0.5 ou 0.35

Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



- A curva produzida é chamada de Curva ROC
- Pois bem, olhando para essa figura, qual é o melhor valor de limiar possível?
- Note que usar 0.2 é melhor do que usar 0 ou 0.05
- Note que usar 0.75 é melhor do que usar 0.5 ou 0.35
- Agora, entre 0.2 e 0.75, depende do que você quer. Você aceita aumentar sua taxa de VP mesmo que isso signifique também aumentar sua taxa de FP?

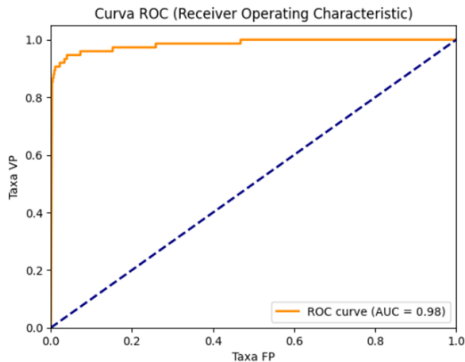
Valor de limiar	Taxa de <i>VP</i>	Taxa de <i>FP</i>
0	100%	100%
0,05	100%	75%
0,2	100%	50%
0,35	75%	50%
0,5	75%	25%
0,75	75%	0%
0,9	50%	0%
0,95	25%	0%
0,95	0%	0%



- A métrica dada pela área sob a curva ROC é chamada de AUC.
- Nesse caso, os valores foram calculados a partir de um modelo de regressão logística. Vamos supor que nós apliquemos também redes neurais ao problema, e que ela nos retorne uma AUC de 0.9. A ideia é dizer que a rede neural performou melhor que a regressão logística.

Voltando ao exemplo de transações financeiras fraudulentas...

	Valor limiar	Taxa de VP	taxa de FP
0	0.00	1.000000	1.000000
1	0.02	0.946667	0.040748
2	0.04	0.933333	0.034525
3	0.06	0.920000	0.031080
4	0.08	0.920000	0.028197
5	0.10	0.920000	0.026070
6	0.12	0.920000	0.024523
7	0.14	0.920000	0.023398
8	0.16	0.920000	0.022361
9	0.18	0.906667	0.021183
10	0.20	0.906667	0.020339
11	0.22	0.906667	0.019759
12	0.24	0.906667	0.019126
13	0.26	0.906667	0.018458
14	0.28	0.906667	0.018054
15	0.30	0.906667	0.017526
16	0.32	0.906667	0.017175
17	0.34	0.906667	0.016753
18	0.36	0.906667	0.016366
19	0.38	0.906667	0.016243
20	0.40	0.906667	0.016014
21	0.42	0.906667	0.015698
22	0.44	0.906667	0.015188
23	0.46	0.906667	0.014854
24	0.48	0.906667	0.014468
25	0.50	0.906667	0.014257
26	0.52	0.906667	0.013975
27	0.54	0.906667	0.013712
28	0.56	0.906667	0.013466
29	0.58	0.906667	0.013290
30	0.60	0.906667	0.013096
31	0.62	0.906667	0.012780
32	0.64	0.906667	0.012516
33	0.66	0.906667	0.012323
34	0.68	0.906667	0.012182
35	0.70	0.906667	0.011989
36	0.72	0.906667	0.011813
37	0.74	0.906667	0.011655
38	0.76	0.906667	0.011532
39	0.78	0.906667	0.011426
40	0.80	0.906667	0.011198
41	0.82	0.906667	0.010969
42	0.84	0.893333	0.010776
43	0.86	0.893333	0.010424
44	0.88	0.893333	0.010126
45	0.90	0.893333	0.009879
46	0.92	0.893333	0.009633
47	0.94	0.893333	0.009317
48	0.96	0.893333	0.008772
49	0.98	0.880000	0.008016
50	1.00	0.853333	0.002953



Observações finais

Observações finais

- As métricas ROC e AUC são muito utilizadas por diversos usuários de ML. Entretanto, pessoalmente falando, eu prefiro usar diretamente a tabela para escolher o valor de limiar.

Observações finais

- As métricas ROC e AUC são muito utilizadas por diversos usuários de ML. Entretanto, pessoalmente falando, eu prefiro usar diretamente a tabela para escolher o valor de limiar.
- Note que as taxas de VP e FP (em função de diferentes valores de limiar) tendo-se como base um único modelo.

Observações finais

- As métricas ROC e AUC são muito utilizadas por diversos usuários de ML. Entretanto, pessoalmente falando, eu prefiro usar diretamente a tabela para escolher o valor de limiar.
- Note que as taxas de VP e FP (em função de diferentes valores de limiar) tendo-se como base um único modelo.
- O ideal é calcular tais métricas com base nos dados de validação, para que as conclusões sejam mais representativas.

Observações finais

- As métricas ROC e AUC são muito utilizadas por diversos usuários de ML. Entretanto, pessoalmente falando, eu prefiro usar diretamente a tabela para escolher o valor de limiar.
- Note que as taxas de VP e FP (em função de diferentes valores de limiar) tendo-se como base um único modelo.
- O ideal é calcular tais métricas com base nos dados de validação, para que as conclusões sejam mais representativas.
- Ao invés de usar as taxas de VP e FP para compor os eixos y e x, respectivamente, muitas pessoas plotam, por exemplo, precisão x recall. Nesse caso, muda o ponto do gráfico que deve ser considerado o “melhor dos mundos”

De olho no código!

De olho no código!

Vamos agora voltar ao exemplo de transações fraudulentas estudado na aula anterior, e implementaremos o cálculo das métricas ROC e AUC.

Acesse o Python Notebook usando o QR code ou o link abaixo:



https://colab.research.google.com/github/xaximpvp2/master/blob/main/codigo_aula23_classificacao_desbalanceada_ROC_AUC.ipynb

- Link para acesso à base de dados: <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud>
- Apenas uma observação adicional: Também existe um OUTRO exemplo no site do Keras que envolve essa questão de transações financeiras fraudulentas. Ele pode ser encontrado nesse site qui: https://keras.io/examples/timeseries/event_classification_for_payment_card_fraud_detection/

Parte 1

Rode todo o código. Responda às questões nele contidas e complete-o, se necessário.

Parte 2

- 1 Fazendo uma análise cuidadosa de como diferentes valores de limiar afetam as taxas de VP e FP resultantes para os dados de validação deste problema, escolha um valor de limiar adequado. Justifique e detalhe sua escolha final.