

56^a COGEF

Comissão de Gestão Fazendária



Inteligência Artificial na Fiscalização de ICMS

Marcelo de Mesquita Lima
Superintendência de Controle e Fiscalização
Secretaria de Economia do Estado de Goiás

Inteligência Artificial:

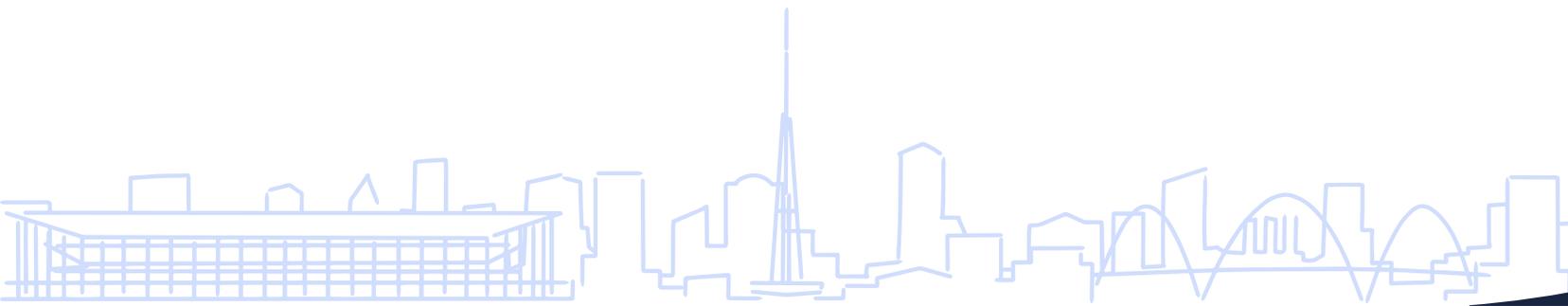
A inteligência artificial (IA) é um campo multidisciplinar da ciência da computação que se concentra no desenvolvimento de sistemas e algoritmos capazes de realizar tarefas que normalmente exigiriam inteligência humana.

Ela tem o potencial de automatizar tarefas rotineiras, auxiliar na tomada de decisões complexas, fornecer insights valiosos a partir de grandes conjuntos de dados e impulsionar a inovação em diversas áreas.



PERGUNTA A SEREM RESPONDIDAS?

- QUANTO VAI CUSTAR O TÁXI?
- QUAL A PROBABILIDADE DE UM CAMINHÃO ESTAR CARREGANDO MAIS PESO DO QUE O DECLARADO NO MANIFESTO?
- COMO IDENTIFICAR CONTRIBUINTES FANTASMAS QUE ESTÃO ATIVOS NO CADASTRO DE CONTRIBUINTES?



Descrição do Problema:

Prever quanto vai custar uma corrida de táxi.

Y = Preço da corrida(R\$)	X1 = Distância(km)	X2 = Hora parada(h)
7,99	1	0
60,99	2	2
38,99	3	1
23,24	4	0,25
47,49	10	0,5
???	15	0,25



DECRETA:

Art. 1º O [Decreto nº 2.096, de 11 de agosto de 2015](#), passa a vigorar com a seguinte alteração:

“Art. 1º

.....

I - R\$ 4,99 (quatro reais e noventa e nove centavos) por bandeirada;

II - R\$ 3,00 (três reais) por quilômetro rodado na bandeira 1;

III - R\$ 3,75 (três reais e setenta e cinco centavos) por quilômetro rodado na bandeira 2;

IV - R\$ 25,00 (vinte e cinco reais) por hora parada;

V - R\$ 2,40 (dois reais e quarenta centavos) por volume adicional transportado, assegurado ao usuário o transporte gratuito de uma mala e dois volumes de mão.” (NR)



Solução do problema do táxi:

Y = Preço da corrida(R\$)	X1 = Distância(km)	X2 = Hora parada(h)
7,99	1	0
60,99	2	2
38,99	3	1
23,24	4	0,25
47,49	10	0,5
???	15	0,25

$$Y = 4,99 + 3,00 * \text{Distância} + 25,00 * \text{hora parada}$$

$$Y = a + b1 * x1 + b2 * x2$$

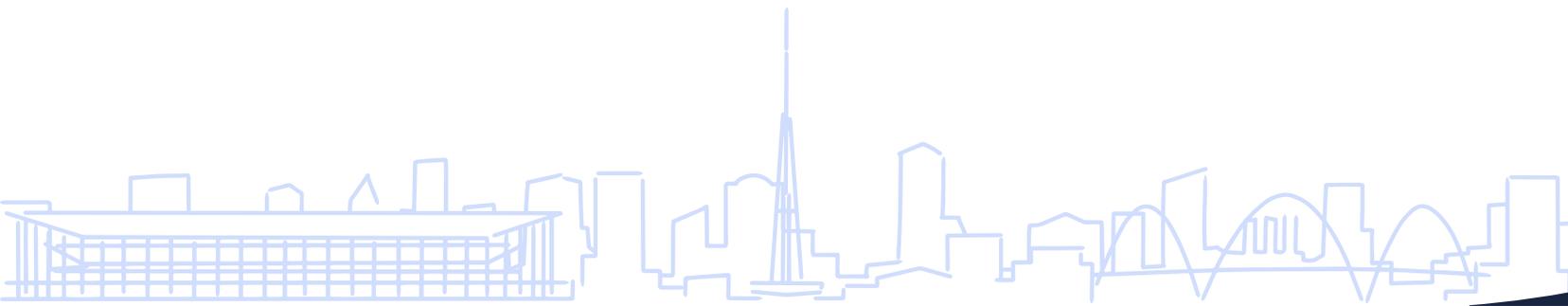


Descrição do Problema:

Estimar a probabilidade de um caminhão estar carregando mais peso que o declarado no manifesto.

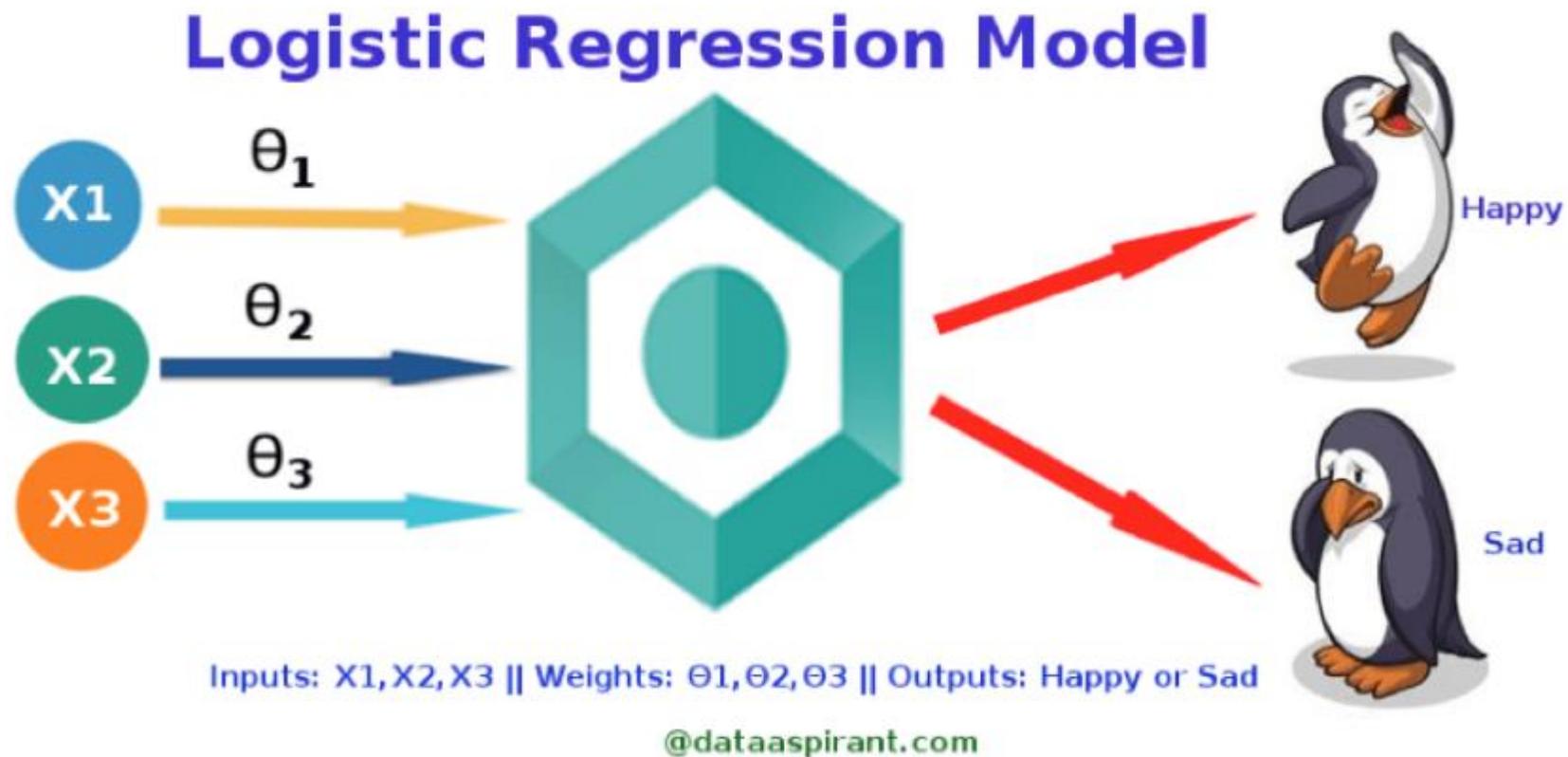
Variáveis explicativas:

- Carga máxima calculada para o mesmo veículo (média das 5 maiores cargas históricas)
- Peso da carga declarada no MDF-e

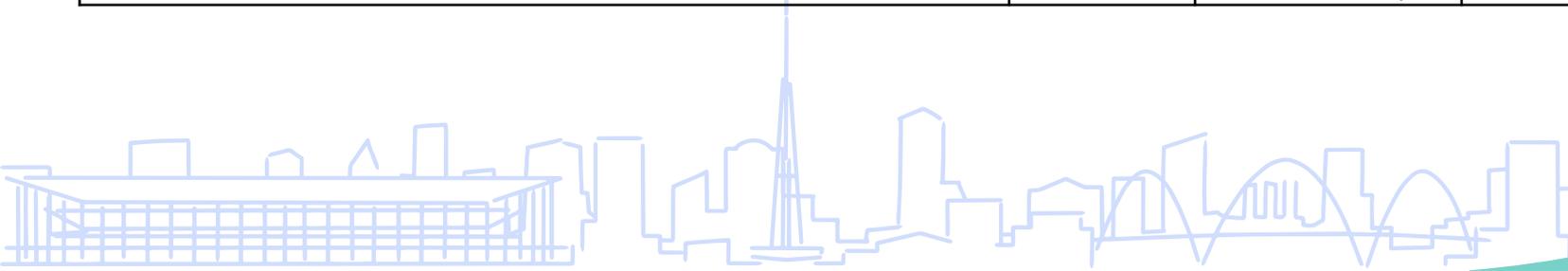


Descrição da Solução:

Utilização de um modelo de regressão logística.



NUMR_CHAVE_ACESSO_MDFE	PLACA	PESO_BRUTO	Carga_Máxima	Mercadoria	Sobrepeso
52230302199856001720580100000634061000932406	GRD6I39	37.480,00	38.022,50	Soja	0
52230302199856001720580100000634071000932438	OMG0E28	38.290,00	38.684,29	Soja	0
52230302199856001720580100000634081000932451	AJB1E28	32.300,00	32.600,00	Soja	0
52230302199856001720580100000634091000932475	NUB2J43	49.420,00	50.830,77	Soja	0
52230302199856001720580100000634101000932492	IAK8038	35.340,00	40.230,00	Soja	0
52230302199856001720580100000634161000932763	CYN9C02	31.760,00	32.212,73	Soja	0
52230302199856001720580100000634171000932809	LYE2271	32.290,00	32.755,00	Soja	0
52230302199856001720580100000634181000932849	JYS5563	32.430,00	32.480,00	Soja	0
52230302199856001720580100000634191000932870	AQE5380	37.560,00	37.794,12	Soja	0
52230302199856001720580100000634201000932910	AAF0C01	31.400,00	31.722,86	Soja	0
52230302199856001720580100000634211000932976	DBC4C53	32.160,00	32.298,75	Soja	0
52230302199856001720580100000634221000932981	PRM5838	50.240,00	50.381,25	Soja	0
52230302199856001720580100000634251000933041	HZE9D93	31.020,00	32.238,00	Soja	0
52230302726560000153580010001448701544211209	HRO7971	37.000,00	47.070,00	Soja	1



Solução do Problema:

$$y = 5,428 + (-8,888 * 10^{-4}) * \text{Peso declarado} + (6,286 * 10^{-4}) * \text{Carga máxima}$$

$$p = \frac{1}{1 + e^{-y}}$$

Variáveis explicativas:

- Carga máxima calculada para o mesmo veículo (média das 5 maiores cargas históricas)
- Peso da carga declarada no MDF-e



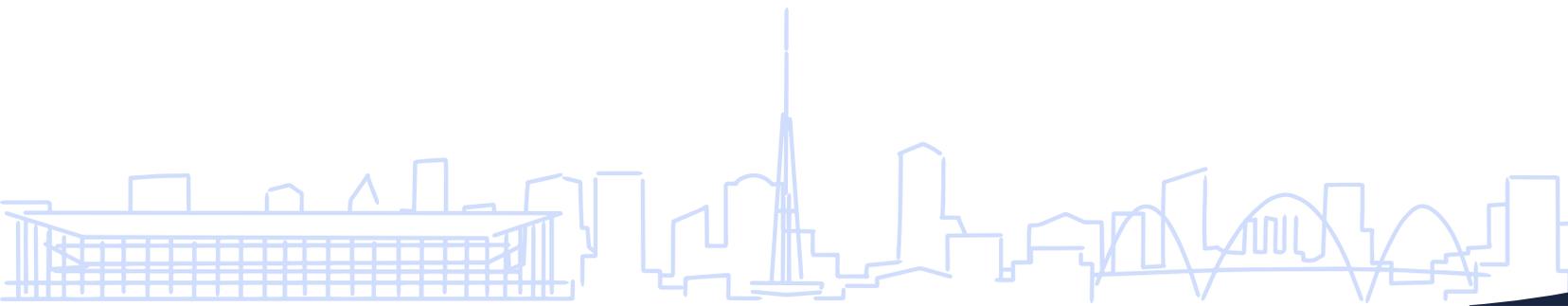
PLACA	PESO_BRUTO	Carga_Máxima	Mercadoria	Sobrepeso	y	p
GRD6I39	37.480,00	38.022,50	Soja		0-3,98	0,02
OMG0E28	38.290,00	38.684,29	Soja		0-4,29	0,01
AJB1E28	32.300,00	32.600,00	Soja		0-2,79	0,06
NUB2J43	49.420,00	50.830,77	Soja		0-6,54	0,00
IAK8038	35.340,00	40.230,00	Soja		0-0,69	0,33
CYN9C02	31.760,00	32.212,73	Soja		0-2,55	0,07
LYE2271	32.290,00	32.755,00	Soja		0-2,68	0,06
JYS5563	32.430,00	32.480,00	Soja		0-2,98	0,05
AQE5380	37.560,00	37.794,12	Soja		0-4,20	0,01
AAF0C01	31.400,00	31.722,86	Soja		0-2,54	0,07
DBC4C53	32.160,00	32.298,75	Soja		0-2,85	0,05
PRM5838	50.240,00	50.381,25	Soja		0-7,56	0,00
HZE9D93	31.020,00	32.238,00	Soja		0-1,88	0,13
HRO7971	37.000,00	47.070,00	Soja		1 2,13	0,89



Descrição do Problema:

Identificar se uma empresa de fato existe ou não, baseado nas informações cadastrais e econômicas relacionadas a ela.

O Delegado Fiscal de Luziânia, Ricardo Costa Pinto, decidiu desenvolver um modelo supervisionado a fim de resolver o problema.



Como é o trabalho:

- O estudo possui como objetivo aplicar predição baseada em redes neurais artificiais capaz de identificar padrões típicos de contribuintes desaparecidos.
- O modelo classifica as empresas, a partir do conjunto de 25 (vinte e cinco) variáveis explicativas, como suspensas (desaparecidas) ou ativas.
- Dessa forma é possível determinar com mais assertividade as pessoas jurídicas a serem fiscalizadas in loco, para averiguação/confirmação da fraude praticada.



Descrição do Trabalho:

- 27.709 contribuintes do cadastro
- Diversos ramos de atividade
- Águas Lindas de Goiás, Cidade Ocidental, Cristalina, Luziânia, Novo Gama, Santo Antônio do Descoberto e Valparaíso de Goiás
- Ficaram de fora : MEI e Produtor Rural



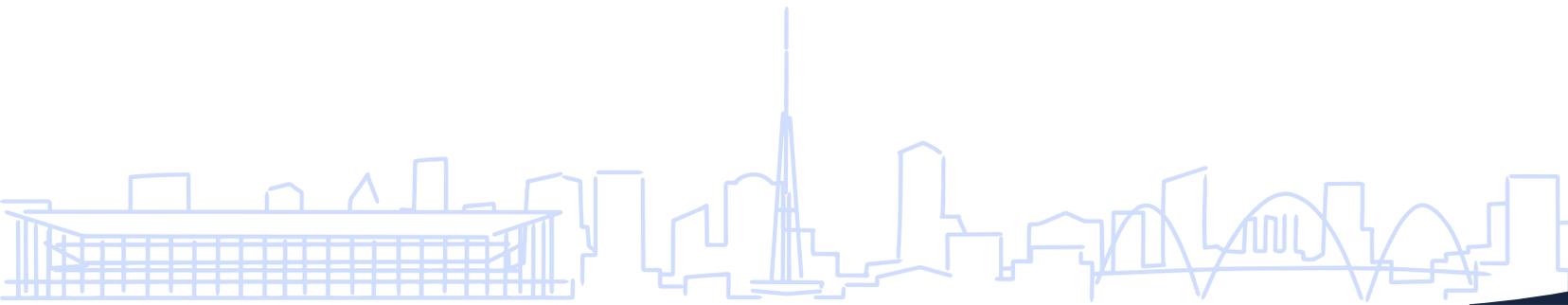
Variáveis explicativas: 25 variáveis

- PORTE
- TIPO ENQUADRAMENTO: Simples/Normal
- ÁREA
- Atividade Econômica
- Indicador de Bloqueio de emissão de notas
- DTE: Sim/Não
- Município
- QTDE_SOCIO_UF_GO e QTDE Sócio outro Estado
- Matriz/Filial
- Valor médio Arrecadação



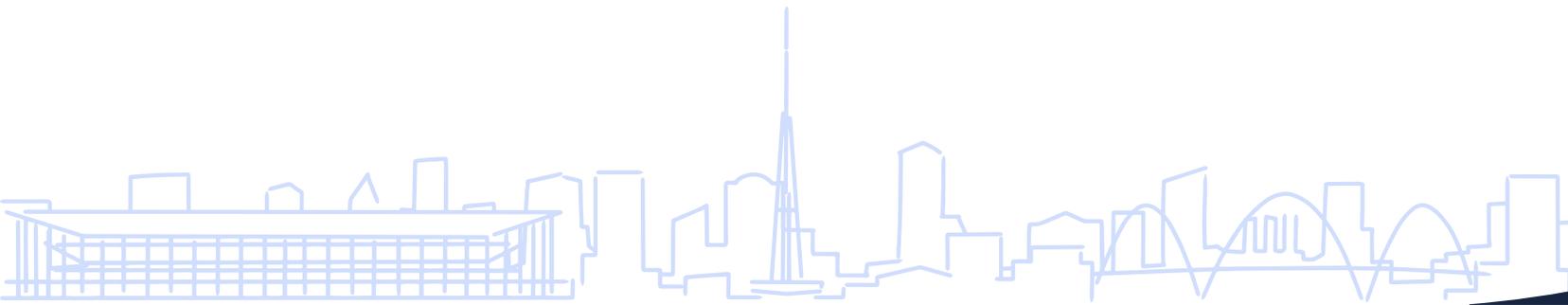
Variáveis explicativas: 25 variáveis

- Valores cartão de Crédito e Débito
- Qtde UF Fornecedor
- Qtde Fornecedor
- Compras: Valor mínimo, máximo e média
- Outras



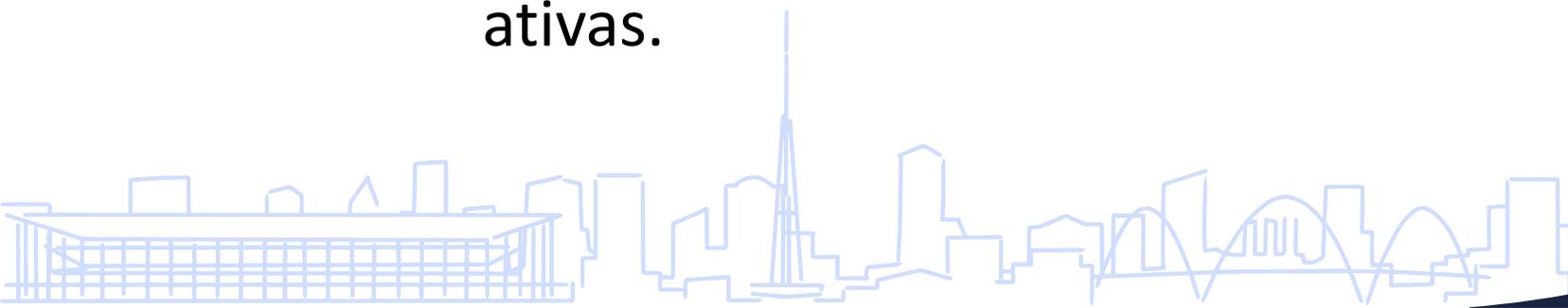
Matriz de confusão:

		Valores Reais	
		Positivo	Negativo
Valores Preditos	Positivo	3.137	1.020
	Negativo	310	3.567



Resultados:

- 3137 empresas foram classificadas como suspensas (desaparecidas), e estavam de fato suspensas;
- 310 empresas foram classificadas como ativas, porém estavam suspensas (falso negativo);
- 3567 empresas foram classificadas como ativas (não suspensas), e a princípio estão ativas.
- 1020 empresas foram classificadas como suspensas (desaparecidas), e a princípio estão ativas.



Indicadores:

- Acurácia (“Accuracy”) – a proporção de observações que foram corretamente previstos, sejam eles verdadeiro positivo ou verdadeiro negativo.
- Sensibilidade (“Sensitivity”) – a proporção de observações positivas que foram identificados corretamente.
- Especificidade (“Specificity”) - a proporção de observações negativas que foram identificados corretamente.



Obrigado!

