



## Introdução

Aplicações e uso

Este material tem como objetivo ser um guia de boas práticas no desenvolvimento e monitoramento de modelos.

Não limitando-se única e exclusivamente a essas métricas, mas é de bom tom que essas sejam ao menos as avaliações básicas para a subida de um modelo em produção levando em consideração o impacto negativo para o negócio que um modelo ruim ou com uma **seleção adversa** pode causar.

De nada vale o conhecimento se for guardado a sete chaves! Conhecimento deve ser compartilhado e dividido, pois guardado é igual dinheiro em um cofre, pois perde o valor e não rende nada.

Sendo assim, caso tenha algo que possa enriquecer ainda mais este material, por favor, não limite-se a guarda-lo somente para você!

Robson Costa

## Introdução

Aplicações e uso

O "Cientista de Dados" ingênuo, em sua tão esperada entrevista de emprego para ingressar no mundo do crédito:

"Modelar risco de crédito é simples, só modelinho de classificação binária"...

Ele se vendeu bem e foi contratado!

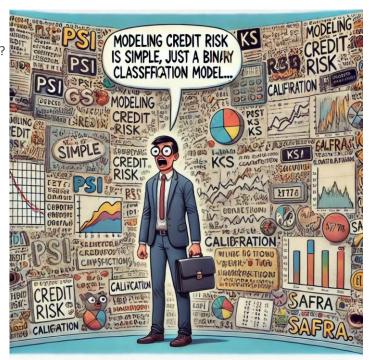
Mas no 1º mês de trabalho... 63



- Como faço a seleção do público?
- Qual a definição da target?
- Entendi, então tem que prever a probabilidade de atrasar mais que X dias em um horizonte de Y meses? Então como constrói a tabela de desenvolvimento mesmo?
- Como assim "backtest"?
- Performance inferida de negados?
- As fontes de dados estão todas na mão, né?
- Safra? Achei que era um termo do agro!
- Tem que cruzar por ID e safra? Como assim em m-1?
- Eita, os dados são desbalanceados?
- Puts, a amostra tem viés de seleção?
- Testes de segmentação e avaliar model fairness?
- Como assim estabilidade? Não é só medir o AUC na amostra de validação?
- OOT?
- Calibração? Pontos de corte e grupos homogêneos de risco?
- PSI? Nunca ouvi falar...
- KS? Lembro que vi na universidade, mas nunca mais usei...
- Swap in? Swap out?

e a lista continua... 🕮







Kolmogorov-Smirnov Index (KS)

O KS identifica o ponto onde ocorre a maior diferenciação entre "bons" e "maus" e é responsável por mensurar a força de discriminação do modelo. Quanto mais o modelo diferencia "bom" de "mau" maior será o KS obtido. Logo, o valor ideal esperado deve ser o mais próximo de 1 (um) possível.

#### **Fórmula**

$$M\acute{a}x(|\% Acum (b) - \% Acum (m)|)$$

#### Exemplo de cálculo

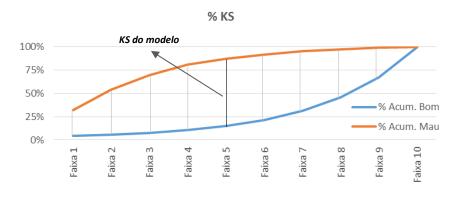
% Bons	% Maus	% Acum. Bom (A)	% Acum. Mau (B)
2,5%	18,1%	2,5%	18,1%
3,0%	16,5%	5,6%	34,6%
3,6%	12,3%	9,2%	46,9%
5,2%	9,3%	14,4%	56,2%
7,5%	8,9%	21,9%	65,1%
9,2%	8,4%	31,1%	73,5%
12,4%	7,8%	43,5%	81,3%
15,6%	7,1%	59,1%	88,4%
19,1%	6,2%	78,2%	94,6%
21,8%	5,4%	100,0%	100,0%
100,0%	100,0%		

Cálcul	o do KS
C =  (B - A)	MÁX(C)
0,16	
0,29	
0,38	
0,42	
0,43	0,43
0,42	
0,38	
0,29	
0,16	
0,00	
KS:	43,2%

#### **Parâmetros**

- De 0% até 10% de variação entre KS atual e desenv. ou diferença até de 5 p.p.
- O De 11% até 20% de variação entre KS atual e desenv. ou diferença até de 10 p.p.
- Acima de 20% de variação entre KS atual e desenv. ou diferença acima de 10 p.p.

#### Representação gráfica





Population Stability Index (PSI)

Esse indicador procura identificar diferenças populacionais entre as distribuições do desenvolvimento do modelo com a atual.

Quanto mais diferentes forem as populações em análise, maior será o PSI obtido. Logo, o valor de PSI esperado deve ser o mais próximo de 0 (zero) possível.

#### **Fórmula**

$$\sum\nolimits_{k=0}^{n} \left(D(s) - O(s)\right) \times \ln \frac{D(s)}{O(s)}$$

#### Exemplo de cálculo

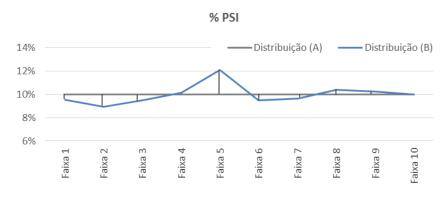
Distribuição (A)	Distribuição (B)
10,0%	9,6%
10,0%	8,9%
10,0%	9,5%
10,0%	10,2%
10,0%	12,1%
10,0%	9,5%
10,0%	9,6%
10,0%	10,4%
10,0%	10,2%
10,0%	10,0%
100,0%	100,0%

	Cálculo do PSI												
C = (B - A)	D = LN(B / A)	C * D											
0,00	-0,05	0,0%											
-0,01	-0,11	0,1%											
-0,01	-0,05	0,0%											
0,00	0,02	0,0%											
0,02	0,19	0,4%											
-0,01	-0,05	0,0%											
0,00	-0,04	0,0%											
0,00	0,04	0,0%											
0,00	0,02	0,0%											
0,00	0,00	0,0%											
	PSI:	0,6%											

#### **Parâmetros**

- De 0% até 10% de variação entre as distribuições
- De 11% até 25% de variação entre as distribuições
- Acima de 25% de variação entre as distribuições

#### Representação gráfica



http://ucanalytics.com/blogs/population-stability-index-psi-banking-case-study/





Similar ao PSI, o VDI procura identificar diferenças entre as distribuições do desenvolvimento do modelo com a safra atual para cada uma das variáveis dos modelos.

Avaliando o VDI para o primeiro segmento, observa-se que duas das variáveis utilizadas possuem diferença em sua distribuição. Apesar da diferença ser positiva, pois reduziu-se a quantidade de *missing* nas variáveis, essa diferença impacta na importância da variável no modelo e por consequência o Beta ao qual está associada.

Idade do Cliente	Baseline (Development + Holdout)			Out-of-Time			Recent		2017.08			
IDADE	#	%	#	%	VDI	#	%	VDI				
01. CLASSE: Missing	26	0.1%	9	0.1%	0.0002	33	0.3%	0.0028		0.0%	0.0036	
02. CLASSE: <= 29.00000	15,435	49.2%	3,688	50.5%	0.0003	5,415	49.0%	0.0000	345	33.0%	0.0649	
03. CLASSE: > 29.00000 AND <= 33.00000	3,351	10.7%	767	10.5%	0.0000	1,162	10.5%	0.0000	100	9.6%	0.0012	
04. CLASSE: > 33.00000 AND <= 37.00000	3,091	9.9%	678	9.3%	0.0003	1,108	10.0%	0.0000	112	10.7%	0.0007	
05. CLASSE: > 37.00000 AND <= 46.00000	4,737	15.1%	1,063	14.6%	0.0002	1,636	14.8%	0.0001	188	18.0%	0.0050	
06. CLASSE: > 46.00000 AND <= 57.00000	3,173	10.1%	711	9.7%	0.0002	1,090	9.9%	0.0001	141	13.5%	0.0097	
07. CLASSE: > 57.00000	1,534	4.9%	388	5.3%	0.0003	604	5.5%	0.0006	159	15.2%	0.1171	
Total	31,347		7,304		0.0016	11,048		0.0036	1,045		0.2022	
Score de Presença Online	Baseline (Dev Hold		Out-of-Time				Recent		2017.08			
EShopper	#	%	#	%	VDI	#	%	VDI				
01. CLASSE: Missing	2,253	7.2%	700	9.6%	0.0069	1,177	10.7%	0.0136	4	0.4%	0.1996	
02. CLASSE: A, C, B, D	460	1.5%	87	1.2%	0.0006	145	1.3%	0.0002	17	1.6%	0.0002	
03. CLASSE: H	26,226	83.7%	6,034	82.6%	0.0001	9,008	81.5%	0.0005	947	90.6%	0.0056	
04. CLASSE: E, F, G	2,408	7.7%	483	6.6%	0.0016	718	6.5%	0.0020	77	7.4%	0.0001	
Total	31,347		7,304		0.0092	11,048		0.0163	1,045		0.2054	



Weight of Evidence (WoE)

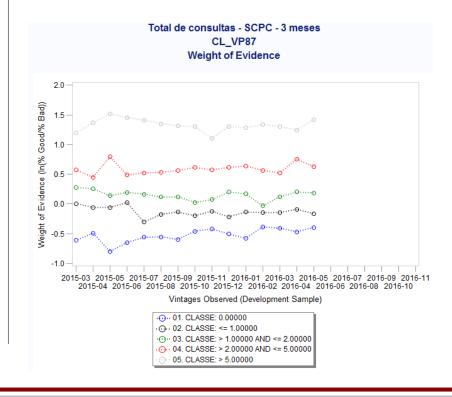
O indicador WoE é o logaritmo natural do % Bons dividido pelo % Maus, utilizado para medir se uma determinada classe está associada com um nível mais elevado ou mais baixo de risco. Se o WoE é positivo, isso significa que há uma proporção maior de Bons do que Maus caindo nessa classe em particular (isto é, menor risco). Se o WoE é negativo, há uma maior proporção de Maus do que Bons (isto é, maior risco). Em geral, os atributos com WoE inferiores receberem uma pontuação mais baixa do que aqueles com maior WoE.

**Fórmula** 

$$(WoE) = \ln(\% Good / \% Bad)$$

Espera-se que o WoE ao longo do tempo para cada uma das classes seja estável e que não haja intersecção, tendo-se assim um poder de discriminação maior para a variável em questão e um modelo estável, ou seja, que se deteriorará com mais tempo em produção.

#### Representação Gráfica Ideal





Coeficiente de Determinação ( $R^2$ )

O  $\mathbb{R}^2$  é uma medida de ajuste de modelos estatísticos em relação aos seus parâmetros e indica o quanto o modelo explica, percentualmente, o evento observado. Ou seja, se o  $\mathbb{R}^2$  for de 0.85, significa que 85% do evento consegue ser explicado pelo modelo.

Todavia, se uma ou mais variáveis do modelo possuírem mais que 60% de correlação (multicolinearidade), o  $\mathbb{R}^2$  será ilusoriamente bom, mas o modelo apresentará elevados níveis de erros padrão.

Espera-se que não haja multicolinearidade entre as variáveis do modelo, obtendo um poder de discriminação maior e um modelo estável, para isso dois testes são realizados:

#### Matriz de Correlações:

Correlação perfeita entre duas variáveis independentes existe se o coeficiente de correlação de Pearson (e / ou de Spearman) for igual a 1 ou -1. Destaca-se onde o coeficiente de correlação de Pearson (e / ou de Spearman) for maior ou igual a 0,6 ou menor ou igual a -0,6, pois este irá revelar qualquer correlação linear considerada forte. Como as variáveis fortemente correlacionadas podem causar instabilidade no procedimento de regressão, recomenda-se que se a correlação é maior que | 0,60 |, uma das variáveis do par altamente correlacionado **deve ser removida**.

#### VIF:

Outra medida para detectar Multicolinearidade é fator de inflação da variância (VIF). Matematicamente esta é calculada como  $1/(1-R^2)$ . Considerase igual a 2 como um limite superior conservador. Se VIF é maior do que ou igual a 2, variáveis "ofensivas" **devem ser removidas**, a menos que haja uma justificação para mantê-las.



Capacidade de Separação

Para administrar o risco de crédito, a modelagem estatística é uma metodologia eficiente baseada em uma amostra de dados já existente com performance conhecida, permitindo a partir dela, criar equações que auxiliam na tomada de decisão para novas populações. Assim, para esse modelo, foi utilizada a regressão logística devido a:

- Técnica eficiente para o desenvolvimento de modelos com respostas binárias (performance bom / mau);
- Consideração de correlação entre as variáveis, identificando relações que não seriam visíveis e eliminando variáveis redundantes;
- Consideração dos efeitos combinados de vários fatores de risco.

Nesta técnica, os clientes da base recebem pesos (Betas) de acordo com a categoria em que eles se enquadram para cada uma das variáveis explicativas selecionadas (x), resultando nas suas probabilidades de serem inadimplentes.

Uma tabela de ganhos é analisada para garantir a <u>ordenação da taxa de</u> <u>maus</u> pelas faixas de score do modelo, ou seja, <u>maior taxa de mau nas piores</u> <u>faixas de score e menor taxa nas melhores faixas</u>.

Espera-se que um modelo bem ajustado não possua inversões em sua ordenação de modo que a classificação de risco tenha consistência.

#### Representação Gráfica Ideal

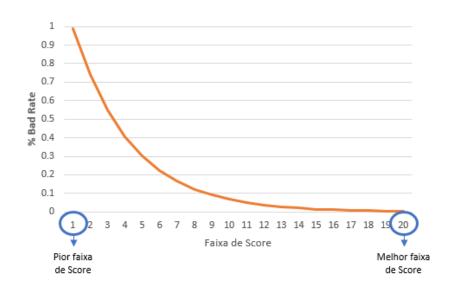




Tabela de Ganhos (Gains Table)

A Tabela de Ganhos é uma ferramenta comumente utilizada na análise de modelos de crédito, em particular no contexto de Modelagem de Risco de Crédito. Ela é utilizada para avaliar o desempenho do modelo, medindo sua capacidade de classificar corretamente os riscos de crédito.

Sua construção é feita ordenando as observações do conjunto de dados de acordo com uma pontuação (score) gerada pelo modelo. Em seguida, as observações são agrupadas em intervalos (também chamados de decis ou percentis) com base nessa pontuação. Para cada intervalo, são calculadas algumas métricas, como a taxa de eventos (proporção de observações positivas) e a taxa de não eventos (proporção de observações negativas) presentes no intervalo.

A tabela de ganhos fornece uma visão do poder discriminatório do modelo, ou seja, sua capacidade de separar os bons riscos dos maus riscos. Ao analisar a tabela, é possível observar como as taxas de eventos e não eventos variam de acordo com o intervalo de pontuação.

Compara-se sempre a tabela com as informações do momento do desenvolvimento (Baseline) com a da safra mais recente maturada, de acordo com a performance oficial do modelo.

Idealmente, espera-se que as taxas de eventos sejam maiores nos intervalos de pontuação mais altos, indicando que o modelo está classificando corretamente os riscos mais elevados como sendo de maior probabilidade. Além disso, espera-se que os resultados destacados na tabela da Baseline e a da safra de referência apresentem comportamentos semelhantes e estejam de acordo com as conclusões sinalizadas em cada um dos indicadores que a compõem.

#### Representação Ideal

						Ga	iin's Table	: Baseline	•							
(#) Bons	(#) Maus	(#) Total	(%) Bons	(%) Maus	(%) Total	(#) Acum. Bons	(#) Acum. Maus	(#) Acum. Total	(%) Acum. Bons	(%) Acum. Maus	(%) Acum. Total	(%) Tx. Maus (Interv.)	Odds	WoE	KS	D
1.653	1.781	3.434	6,64%	17,33%	9,76%	1.653	1.781	3.434	6,64%	17,33%	9,76%	51,86%	0,93	-0,96	10,69%	10,
1.975	1.628	3.603	7,93%	15,84%	10,24%	3.628	3.409	7.037	14,57%	33,17%	20,00%	45,18%	1,21	-0,69	18,61%	5,
2.249	1.513	3.762	9,03%	14,72%	10,69%	5.877	4.922	10.799	23,59%	47,89%	30,69%	40,22%	1,49	-0,49	24,30%	2,
2.170	1.114	3.284	8,71%	10,84%	9,33%	8.047	6.036	14.083	32,31%	58,73%	40,03%	33,92%	1,95	-0,22	26,43%	0,
2.631	1.288	3.919	10,56%	12,53%	11,14%	10.678	7.324	18.002	42,87%	71,27%	51, 16%	32,87%	2,04	-0,17	28,40%	0,
2.282	903	3.185	9,16%	8,79%	9,05%	12.960	8.227	21.187	52,03%	80,05%	60,22%	28,35%	2,53	0,04	28,02%	0,
2.996	883	3.879	12,03%	8,59%	11,02%	15.956	9.110	25.066	64,06%	88,64%	71,24%	22,76%	3,39	0,34	24,58%	1,1
2.637	535	3.172	10,59%	5,21%	9,02%	18.593	9.645	28.238	74,65%	93,85%	80, 26%	16,87%	4,98	0,71	19,20%	3,8
1.933	270	2.203	7,76%	2,63%	6,26%	20.526	9.915	30.441	82,41%	96,48%	86,52%	12,26%	7,16	1,08	14,07%	5,5
4.382	362	4.744	17,59%	3,52%	13,48%	24.908	10.277	35.185	100,00%	100,00%	100,00%	7,63%	12, 10	1,61	0,00%	22,
	1.653 1.975 2.249 2.170 2.631 2.282 2.996 2.637 1.933	1.653 1.781 1.975 1.628 2.249 1.513 2.170 1.114 2.631 1.288 2.282 903 2.996 883 2.637 535 1.933 270	1.653 1.781 3.434 1.975 1.628 3.603 2.249 1.513 3.762 2.170 1.114 3.284 2.631 1.288 3.919 2.282 903 3.185 2.996 883 3.879 2.637 535 3.172 1.933 270 2.203	1.653 1.781 3.434 6.64% 1.975 1.628 3.603 7.93% 2.249 1.513 3.762 9.03% 2.170 1.114 3.284 8.71% 2.631 1.288 3.919 10.56% 2.282 903 3.185 9.16% 2.996 883 3.879 12.03% 2.996 883 3.879 12.03% 2.997 2.997 2.003 7.76%	1.653 1.781 3.434 6.64% 17.33% 15.84% 22.49 1.513 3.762 9.03% 14.72% 2.470 1.114 3.284 8.71% 10.84% 2.631 1.288 3.919 10.56% 12.53% 2.996 883 3.879 12.03% 8.59% 2.996 883 3.879 12.03% 8.59% 2.995 8.33 3.79 12.03% 8.59% 2.995 2.997 2.03% 7.76% 2.637 1.933 2.70 2.203 7.76% 2.637	1.653 1.781 3.434 6.64% 17.33% 9.76% 19.75 1.628 3.603 7.93% 15.84% 10.24% 22.49 1.513 3.762 9.03% 14.72% 10.69% 2.631 1.288 3.919 10.56% 12.53% 11.14% 2.282 9.03 3.185 9.16% 8.75% 9.05% 2.996 883 3.879 12.03% 8.5% 11.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.35 3.172 10.59% 5.21% 9.02% 2.637 5.63 2.63 2.63 2.63 2.63 2.63 2.63 2.63 2	Bons	1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   1.975   1.628   3.603   7.93%   15.84%   10.24%   3.628   3.409   2.249   1.513   3.762   9.03%   14.72%   10.69%   5.877   4.922   2.170   1.114   3.284   8.71%   10.81%   9.33%   8.047   6.036   2.631   1.288   3.919   10.56%   12.53%   11.14%   10.678   7.324   2.282   903   3.185   9.16%   8.79%   9.05%   12.960   8.227   2.996   883   3.879   12.03%   8.99%   11.02%   15.956   9.110   2.637   5.35   3.172   10.59%   5.21%   9.02%   18.593   9.645   1.933   2.70   2.203   7.76%   2.63%   6.26%   2.0526   9.915	1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   1.781   3.784   9.76%   1.653   1.781   3.434   1.975   1.628   3.603   7.93%   15.84%   10.24%   3.628   3.409   7.037   2.249   1.513   3.762   9.03%   14.72%   10.69%   5.877   4.922   10.799   2.170   1.114   3.284   8.71%   10.84%   9.33%   8.047   6.036   14.083   2.631   1.288   3.919   10.56%   12.53%   11.14%   10.678   7.324   18.002   2.282   903   3.185   9.16%   8.79%   9.05%   12.960   8.227   2.187   2.296   883   3.879   12.03%   8.99%   11.02%   15.966   9.110   2.5066   2.637   5.35   3.172   10.59%   5.21%   9.02%   18.593   9.645   28.238   1.933   2.70   2.203   7.76%   2.63%   6.26%   20.526   9.915   30.441   3.348   3.488   3.4	1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   1.975   1.628   3.603   7.93%   15.84%   10.24%   3.628   3.409   7.037   14.57%   12.249   1.513   3.762   9.03%   14.72%   10.69%   5.877   4.922   10.799   22.59%   2.170   1.114   3.284   8.71%   10.84%   9.33%   8.047   6.036   14.083   32.31%   2.631   1.288   3.919   10.56%   12.53%   11.14%   10.676   7.324   18.002   42.87%   2.282   903   3.185   9.16%   8.79%   10.05%   12.960   8.227   21.187   52.03%   2.996   883   3.879   12.03%   8.59%   11.02%   15.956   9.110   25.066   64.06%   2.637   5.35   3.172   10.59%   5.21%   9.02%   18.593   9.645   28.238   7.655%   1.933   2.70   2.203   7.76%   2.63%   6.26%   20.526   9.915   30.441   82.41%   32.41%	1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   17.33%   9.76%   1.653   1.781   3.434   6.64%   17.33%   1.975   1.628   3.603   7.93%   15.84%   10.24%   3.628   3.409   7.037   14.57%   33.17%   2.249   1.513   3.762   9.03%   14.72%   10.66%   5.877   4.922   10.799   23.59%   47.89%   2.107   1.114   3.284   8.71%   10.64%   9.33%   8.047   6.036   14.083   32.31%   58.73%   2.631   1.288   3.919   10.56%   12.53%   11.14%   10.678   7.324   18.002   42.87%   71.27%   2.282   903   3.185   9.16%   8.79%   9.05%   12.960   8.227   21.187   52.03%   80.05%   2.282   903   3.185   9.16%   8.59%   11.02%   15.956   9.110   25.066   64.06%   88.64%   2.637   5.35   3.172   10.59%   5.23%   0.02%   18.593   9.645   28.238   74.65%   93.85%   1.933   2.70   2.203   7.76%   2.63%   6.26%   20.526   9.915   30.441   82.41%   66.48%	Bons   Maus   Total   Bons   Bons   Maus   Total   Bons   Bons	Bons   Maus   Total   To	Bons   Maus   Total   Bons   Maus   Total   Bons   Maus   Total   Bons   Maus   Total   (Interv.)	Bons   Maus   Total   Bons   Maus   Total   Bons   Maus   Total   Bons   Maus   Total   Bons   Bons   Maus   Total   Bons   Bons	Bons   Maus   Total   Bons   Total   Total

							-	iii 5 iubte.	E-O-E-E-I-O-I-								
Faixa Score	(#) Bons	(#) Maus	(#) Total	(%) Bons	(%) Maus	(%) Total	(#) Acum. Bons	(#) Acum. Maus	(#) Acum. Total	(%) Acum. Bons	(%) Acum. Maus	(%) Acum. Total	(%) Tx. Maus (Interv.)	Odds	WoE	KS	IV
1,00	1.756	1.968	3.724	2,15%	7,56%	3,46%	1.756	1.968	3.724	2,15%	7,56%	3,46%	52,85%	0,89	-1,26	5,41%	6,81
2,00	3.041	2.608	5.649	3,72%	10,02%	5,24%	4.797	4.576	9.373	5,87%	17,59%	8,70%	46,17%	1, 17	-0,99	11,72%	6,24
3,00	5.077	3.680	8.757	6,21%	14,14%	8,13%	9.874	8.256	18.130	12,08%	31,73%	16,83%	42,02%	1,38	-0,82	19,65%	6,52
4,00	5.398	2.997	8.395	6,60%	11,52%	7,79%	15.272	11.253	26.525	18,69%	43,25%	24,62%	35,70%	1,80	-0,56	24,56%	2,73
5,00	7.843	3.842	11.685	9,60%	14,76%	10,84%	23.115	15.095	38.210	28,28%	58,01%	35,46%	32,88%	2,04	-0,43	29,73%	2,23
6,00	8.118	3.217	11.335	9,93%	12,36%	10,52%	31.233	18.312	49.545	38,21%	70,37%	45,98%	28,38%	2,52	-0,22	32,16%	0,53
7,00	11.999	3.411	15.410	14,68%	13, 11%	14,30%	43.232	21.723	64.955	52,89%	83,48%	60,28%	22,13%	3,52	0,11	30,59%	0,18
8,00	11.068	2.028	13.096	13,54%	7,79%	12,15%	54.300	23.751	78.051	66,44%	91,28%	72,44%	15,49%	5,46	0,55	24,84%	3,18
9,00	8.114	1.077	9.191	9,93%	4,14%	8,53%	62.414	24.828	87.242	76,36%	95,42%	80,96%	11,72%	7,53	0,87	19,05%	5,06
10,00	19.318	1.193	20.511	23,64%	4,58%	19,04%	81.732	26.021	107.753	100,00%	100,00%	100,00%	5,82%	16, 19	1,64	0,00%	31,24



Concentração de Escore

Em um range de 0 a 1000, espera-se que a concentração em pontos específicos de escore não ultrapasse 3%. Visto que isso mostra o quão granular está o escore.

Caso esse evento ocorra é possível que tenham variáveis explicativas que não foram incluídas no modelo.



Contribuição das Variáveis

Espera-se que as variáveis possuam uma contribuição acima de 2%, a não ser que a variável tenha algum sentido para o negócio que justifique sua permanência no modelo.

A métrica é definida como o valor da estatística de Wald da variável dividido pela soma dos Walds, com exceção do intercepto.

Para modelos de Machine Learning existem outras formas de avaliarmos a contribuição/importância das variáveis, como o *Permutation Importance* (PIMP) — O método de permutação de importâncias é "um algoritmo que avalia a relevância de um conjunto de recursos em relação a um vetor de resposta. O algoritmo PIMP permuta n vezes o vetor de resposta. Para cada permutação do vetor de resposta, a relevância para todas as variáveis preditoras é avaliada. Isso leva a um vetor de medidas de importância para cada variável, que chamamos de importâncias nulas. O algoritmo PIMP ajusta uma distribuição de probabilidade à população de importâncias nulas, que o usuário pode escolher entre as seguintes: Gaussiana, lognormal ou gama". (Altmann et al., 2010).



Acurácia

Utilizado para definir a proximidade de um resultado experimental com o seu valor real, em que quanto maior a acurácia, mais autêntico ele é.

Em sua aplicação na tecnologia, ela também pode ser definida como a proximidade de um resultado experimental com o seu valor de referência real. Assim, ela determina o grau de exatidão. Em outras palavras, é uma medida para quantificar o nível de erro entre um valor desejado, ideal, e um valor real ou medido.

Embora possam ter significados parecidos e até mesmo serem utilizados muitas vezes como sinônimos, é muito importante deixar claro que **acurácia e precisão** são conceitos diferentes. Para evitar equívocos, é preciso ter clareza sobre as especificidades de cada um.

A precisão é o grau de variação gerado por diferentes medições. Dessa forma, quanto mais preciso for um processo, menor será a variação entre os valores obtidos. Já a acurácia é uma espécie de soma entre exatidão e precisão. Ou seja, nela, os resultados obtidos por uma experiência não apenas devem ser precisos, mas também precisam estar perto do valor de referência ou real usado como base.

Ao aplicar Inteligência Artificial e Machine Learning em modelos antifraude, por exemplo, é preciso trabalhar com um nível altíssimo de acurácia, para garantir que os consumidores estejam protegidos e que as ações criminosas fiquem sob controle. Em outras palavras, assim é possível ter a certeza de que os métodos utilizados para os processos têm resultados satisfatórios e que a aplicação da tecnologia é feita da maneira mais correta possível.

Existem muitas métricas como Acurácia, Recall, Precision, F1-Score e AUC/ROC, que avaliam seu modelo, antes de colocá-lo em produção. Se você não sabe sobre o que eu estou falando, dá uma olhada nesses artigos e veja o que são essas métricas e quais as diferenças entre elas:

- https://www.flai.com.br/juscudilio/como-calcular-as-metricas-de-validacao-dos-modelos-de-machine-learning-em-python/
- https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html
- https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1 score.html
- https://blogbr.clear.sale/conheca-e-saiba-como-aplicar-a-acuracia

Vale lembrar que quando a amostra é desbalanceada, não é recomendado usar a acurácia. Nesse caso utilizamos a Precision ou Recall, depende do objetivo da análise.



Sentido das variáveis

Com o modelo proposto, cada variável deve ser analisada individualmente e observado o sinal dos Betas e se os mesmos coincidem com a lógica da variável.

Há situações onde uma variável pode entrar com o "sinal trocado" como, por exemplo, em um modelo de seguro veicular é sabido que quanto menor a idade maior é a probabilidade de sinistro. Mas se no modelo a variável estiver penalizando o condutor com idade mediana, significa que algo está errado/contraditório no modelo.

Outro exemplo seria em um modelo de concessão de crédito (*Application Score*), se o cliente **com** restrição/negativação no mercado tivesse uma pontuação melhor que um cliente adimplente.

