



Métricas de avaliação dos modelos de aprendizado de máquina





Modelos de classificação





Métricas - Matriz de confusão



Através da matriz de confusão conseguimos ver os acertos e erros do modelo. É um método importante e que ajuda na estratégia a ser adotada na avaliação do modelo com base nas necessidades do negócio.

		P R E D I T O	
R E A L	POSITIVO	 POSITIVO	 NEGATIVO
	POSITIVO	  TP verdadeiro positivo	  FN falso negativo
	NEGATIVO	  FP falso positivo	  TN verdadeiro negativo





Métricas - Acurácia (Accuracy)

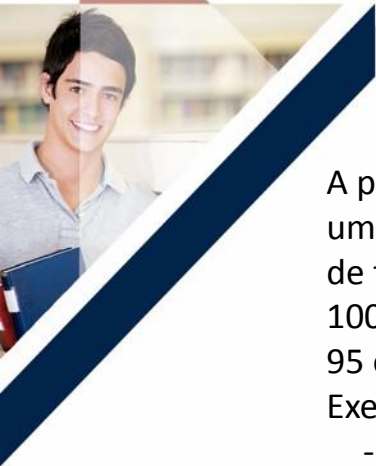


Esta métrica nos apresenta o quão assertivo foi o modelo considerando todos os exemplos classificados corretamente independente da classe. Por exemplo se temos 1000 dados de teste e o modelo acertou 950, temos 95% de acurácia, os outros 50 clientes podem ser falsos positivos ou falsos negativos mas a medida se manteria em 95% levando a crer que o modelo tem uma performance satisfatória.

Apesar de ser simples e fácil de interpretar pode não ser muito usada por não trazer uma representação por classe, em um caso de classes desbalanceadas a acurácia poderá ser boa mas isso não quer dizer que o modelo está bom do ponto de vista de assertividade em classes distintas.

$$\text{Acurácia} = \frac{\checkmark \text{👍 TP} + \checkmark \text{👎 TN}}{\checkmark \text{👍 TP} + \checkmark \text{👎 TN} + \text{✗} \text{👍 FP} + \text{✗} \text{👎 FN}}$$





Métricas - Precisão (Precision)



A precisão é uma métrica que dá mais peso ao falso positivo, Podemos definir a precisão como uma métrica que responde a pergunta: Dos exemplos classificados como positivos, quantos são de fato positivos? Logo, esta métrica dá maior peso aos erros por falso positivo, no exemplo de 1000 exemplos de teste com uma precisão de 95%, diz que a cada 100 amostras de dados apenas 95 estejam classificados de fato corretamente na classe indicada.

Exemplo:

- 90 Casos classificados como positivos
- 67 São positivos
- $67/90 = 74\%$ Precisão

$$\text{Precisão} = \frac{\text{✓👍 TP}}{\text{✓👍 TP} + \text{✗👍 FP}}$$





Métricas - Recall (Sensibilidade)



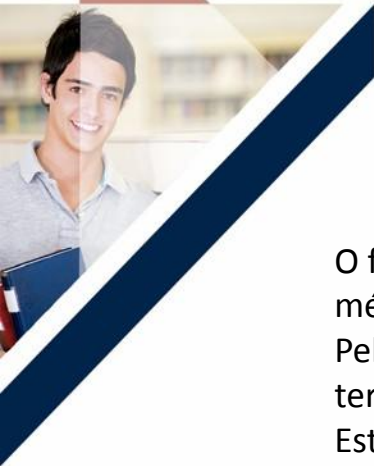
O Recall ao contrário da precisão é uma métrica que pesa mais os erros por falso negativo, poderíamos dizer que ela responde a pergunta: De todos os exemplos que são positivos quantos foram classificados como positivos?

Exemplo:

- 119 Positivos
- 70 Classificados como positivos
- $70/119 = 58\%$ Recall

$$\text{Revocação} = \frac{\checkmark \text{👍 TP}}{\checkmark \text{👍 TP} + \text{✗ 👎 FN}}$$





Métricas - F-Measure (f1 Score)



O f1-score é a média harmônica entre precision e recall, ele se dá pela fórmula abaixo e é uma métrica que pode ser um meio termo para avaliação entre precision e recall .

Pelo fato de usar a média harmônica, se tivermos valores baixos em precision e recall também teremos um f1-score baixo.

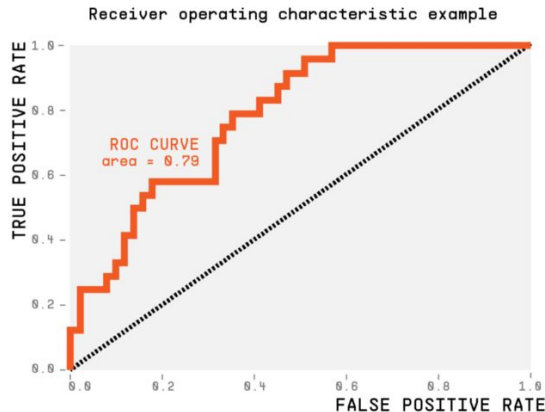
Esta métrica tende a ser uma métrica que representa melhor a assertividade do modelo pois vai ponderar duas métricas que usam visões tanto para FP quanto para FN.

$$F_1 = 2 * \frac{\text{precisão} * \text{recall}}{\text{precisão} + \text{recall}}$$



Métricas - Curva ROC

Dado que atribuímos a probabilidade de um dado pertencer a uma classe positiva, é necessário definir um limiar (threshold) que definirá quando um dado é atribuído a uma classe ou outra. Esse limiar tem relação direta com as métricas que vimos anteriormente, e a definição deste limiar deve considerar o custo de cada erro (FN, FP).



A curva ROC pode ser usado para medir a eficiência de diferentes limiares, ela usa as taxas de falso positivo e verdadeiro positivo. A curva ROC é apresentada através de uma representação gráfica que quanto mais próxima do canto superior esquerdo maior será a cobertura e consequentemente melhor o modelo.

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$



Modelos de regressão



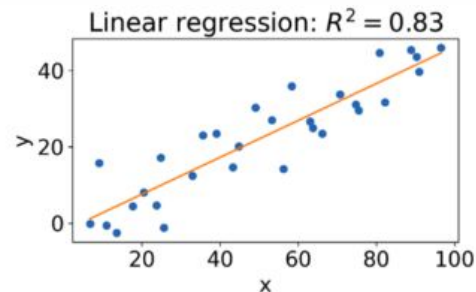
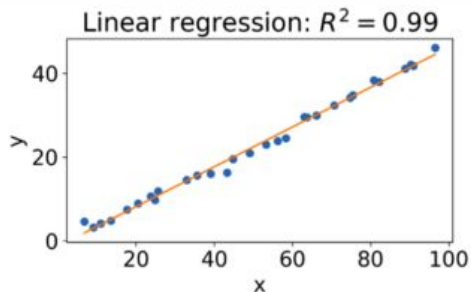
Métricas - R2

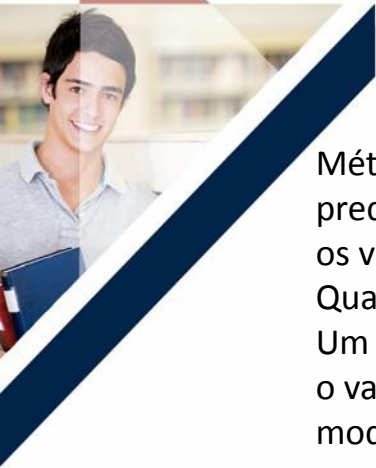
Métrica que expressa a quantidade da variância explicada pelo modelo. Representada por um número que vai de 0 a 1, este valor é o percentual da variância que pode ser prevista pelo modelo de regressão, logo, nos diz o quão nosso modelo se aproxima dos dados reais.

Um R-quadrado de 0,85, nos diz que 85% da variância dos dados pode ser prevista pelo nosso modelo, enquanto que os 15% restantes são a variância residual.

Esta métrica pode ser aplicada com perfeição em casos univariados, ou seja, apenas uma variável independente. Mesmo em casos de overfitting essa métrica continua alta, o R-quadrado é enviesado pelas correlações das variáveis independentes e com o acréscimo de mais features.

$$R^2 = 1 - \frac{\text{Variança Residual}}{\text{Variança Total}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$





Métricas - MSE (mean Squared Error)

Métrica que consiste na média dos erros do modelo ao quadrado. A diferença entre o valor predito pelo modelo e o valor real o valor é elevado ao quadrado, este processo é feito com todos os valores preditos e reais, somamos os valores e dividimos pela quantidade de valores preditos. Quanto maior for o MSE pior a qualidade do modelo.

Um dos pontos de atenção é que um valor predito muito distante do valor real pode elevar demais o valor de MSE, o que torna uma métrica excelente para casos onde a tolerância a erros nos modelos é pequena.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$





Métricas - RMSE (Root Mean Squared Error)

Métrica que consiste na raiz da média dos erros do modelo ao quadrado. Torna a interpretabilidade da métrica melhor mas ainda penaliza bastante valores muito distantes entre o previsto e o real.



$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}$$





Métricas - MAE (Mean Absolute Error)

Métrica que consiste na raiz da média das distâncias entre o valor real e o valor predito. Esta métrica não pune tão severamente outliers do modelo.
ideal para dados sazonais e muitos dados, como tendências.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$





Métricas - MAPE (Mean Absolute Percentual Error)

Métrica que consiste no cálculo percentual de erro entre o valor predito e valor real.

Da mesma forma que o MSE e MAE, quanto menor o valor da métrica melhor seria o modelo.

Por ser uma métrica percentual se torna uma métrica mais simples de interpretar e entender.

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$





Métricas - RMSLE (Root Mean Squared Logarithmic Error)

Esta métrica é similar ao RMSE, mas é usado o logaritmo para evitar que seja penalizado valores com diferenças muito grandes entre o predito e o real.

$$\text{RMSLE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log(y_i + 1) - (\log \hat{y}_i + 1))}$$





Obrigado

