

Questões de Classificação:

Davi Ventura C. Perdigão - 82148

Questão 1:

Uma empresa desenvolveu um modelo de classificação para identificar fraudes em transações financeiras. O modelo foi treinado com uma base de dados de 10.000 transações, sendo 500 fraudulentas e 9.500 legítimas. Após avaliar o modelo com uma base de teste de 1.000 transações, foram obtidos os seguintes resultados:


Verdadeiros Positivos: 200

Falsos Positivos: 50

Verdadeiros Negativos: 700

Falsos Negativos: 50

Calcule a acurácia, precisão, recall e F1 do modelo.

$$\begin{aligned} \text{(1-)} \quad & VP = 200 \quad \rightarrow \text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN} = \frac{200 + 700}{200 + 50 + 700 + 50} = \frac{900}{1000} = 0,9 \text{ (90\%)} \\ & FP = 50 \\ & VN = 700 \\ & FN = 50 \\ \\ & \rightarrow \text{Precisão} = \frac{VP}{VP + FP} = \frac{200}{200 + 50} = \frac{200}{250} = 0,8 \text{ (80\%)} \\ & \rightarrow \text{Recall} = \frac{VP}{VP + FN} = \frac{200}{200 + 50} = \frac{200}{250} = 0,8 \text{ (80\%)} \\ \\ & \rightarrow F1 = \frac{2(\text{precisão} \cdot \text{recall})}{\text{precisão} + \text{recall}} = \frac{2(0,8 \cdot 0,8)}{0,8 + 0,8} = \frac{2 \cdot 0,64}{0,16} = \frac{1,28}{0,16} = 8 \text{ (80\%)} \end{aligned}$$


© & ™ Lucasfilm Ltd.

Questão 2:

Um modelo de classificação foi treinado para identificar se uma pessoa é ou não portadora de uma doença rara. Ao avaliar o modelo com uma base de teste de 200 pessoas, foram obtidos os seguintes resultados:

Verdadeiros Positivos: 150

Falsos Positivos: 20

Verdadeiros Negativos: 25

Falsos Negativos: 5

Calcule a acurácia, precisão, recall e F1 do modelo.

$$\begin{aligned} \text{(2-)} \quad & VP = 150 \quad \rightarrow \text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN} = \frac{150 + 25}{150 + 20 + 25 + 5} = \frac{175}{200} = 0,875 \text{ (87,5\%)} \\ & FP = 20 \\ & VN = 25 \\ & FN = 5 \\ \\ & \rightarrow \text{Precisão} = \frac{VP}{VP + FP} = \frac{150}{150 + 20} = \frac{150}{170} = 0,88 \text{ (88\%)} \\ \\ & \rightarrow \text{Recall} = \frac{VP}{VP + FN} = \frac{150}{150 + 5} = \frac{150}{155} = 0,96 \text{ (96\%)} \\ & \rightarrow F1 = \frac{2(0,88 \cdot 0,96)}{0,88 + 0,96} = \frac{2 \cdot 0,8448}{1,84} = \frac{1,6896}{1,84} = 0,91 \text{ (91\%)} \end{aligned}$$

Questão 3

Um modelo de classificação binária produziu os seguintes resultados: Acurácia = 0,80, Precisão = 0,75 e Recall = 0,85. Qual seria a possível explicação para a discrepância entre a precisão e o recall?

R: Antes de responder diretamente a pergunta, para facilitar a compreensão, segue uma breve explicação sobre as medidas de desempenho descritas acima, bem como respectivos exemplos de aplicação de cada uma delas:

- **Acurácia:** proporção de acertos em relação ao total de classificações realizadas. Na maioria dos casos, a precisão e o recall são mais exatos para avaliar o desempenho de um modelo. Exemplo de aplicação: Um modelo de classificação de doentes rotulou corretamente 92 imagens e errou 8, portanto a acurácia seria de 90%;
- **Precisão:** taxa de acertos de um modelo, calculada com a proporção de acertos em relação ao total de classificações. Sendo assim, a precisão tem como objetivo reduzir falsos positivos. Exemplo de aplicação: Um modelo de classificação de clima/tempo avaliou 1.000 cidades e classificou corretamente sua temperatura, umidade e probabilidade de chuva em 875 delas, a precisão do modelo seria de 87,5%.
- **Recall:** proporção de verdadeiros positivos em relação ao total de resultados positivos do conjunto de dados, buscando reduzir falsos negativos. Ou seja, é a capacidade do modelo de identificar corretamente os resultados positivos na base de dados. Exemplo de aplicação: um modelo é treinado para identificar pães em imagens e é testado em um conjunto que contém 1.000 imagens de pães e 120 imagens de roscas, o recall do modelo é a proporção de imagens de pães que o modelo conseguiu identificar corretamente. Se o modelo identificou corretamente 770 das 1.000 imagens de cães, o recall do modelo é de 0,77 (77%).

Com base nisso, é de se imaginar que a discrepância entre a precisão e o recall em um modelo pode ocorrer por diferentes motivos, dependendo do contexto em que o modelo está inserido. Um dos motivos pode ocorrer se o conjunto de dados usado para treinar o modelo contém uma grande desigualdade, com uma classe sendo muito mais comum do que outra. Por exemplo, uma classe "doente" pode ser muito menos comum do que a classe "saudável". Se o modelo for treinado com um conjunto de dados desbalanceado, pode ter um bom desempenho na classe majoritária (saudável), mas ter um desempenho ruim na classe minoritária (doente), o que levaria a uma baixa recall e alta precisão. Por outro motivo, pode ocorrer justamente o contrário (precisão baixa, mas a recall alta) que seria a presença de erros de classificação. Por exemplo, se o modelo classificar muitos exemplos como verdadeiros negativos quando eles na verdade são falsos negativos.

Questão 4

Suponha que temos um modelo de classificação com acurácia de 0,90, precisão de 0,85 e recall de 0,95. Calcule o valor de F1-score para esse modelo e interprete o resultado.

$$\textcircled{4} \quad F1\text{-score} = \frac{2(\text{precisão} \cdot \text{recall})}{\text{precisão} + \text{recall}}$$
$$\frac{2(0,85 \cdot 0,95)}{0,85 + 0,95} = \frac{2 \cdot 0,80}{1,80} = \frac{1,6}{1,8} = 0,88$$

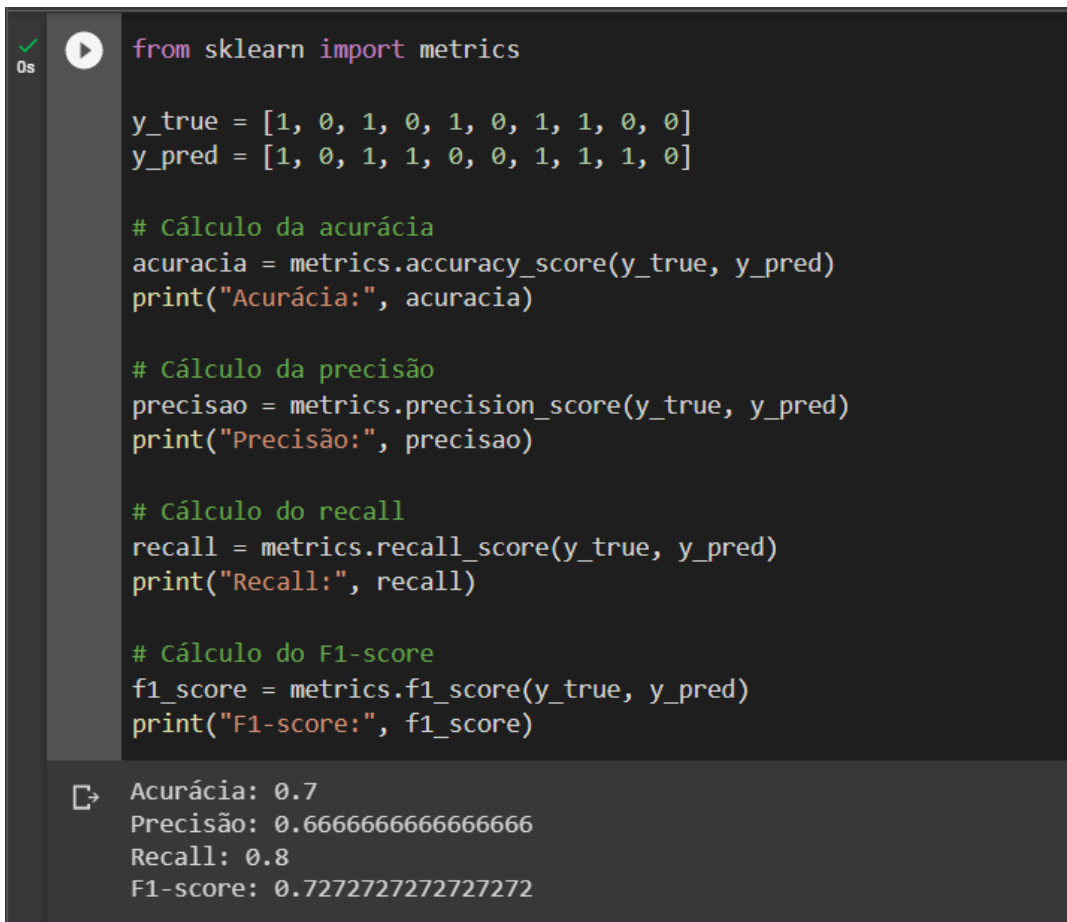
R: O F1-score indica uma medida agregada de precisão e recall do modelo. Sendo assim, o resultado 0,88 indica que o modelo tem um desempenho equilibrado em relação à precisão e recall, mas está mais próximo de uma boa precisão (0,85) do que de um bom recall (0,95). Especificamente, isso significa que o modelo é capaz de identificar corretamente cerca de 88% dos casos verdadeiros positivos, ao mesmo tempo em que minimiza os falsos positivos e os falsos negativos. No geral, um F1-score de 0,88 é considerado um bom resultado e indica que o modelo é confiável na classificação dos dados.

Questão 5:

Considere um modelo de classificação binária que foi treinado e testado com as seguintes previsões:

```
y_true = [1, 0, 1, 0, 1, 0, 1, 1, 0, 0]
y_pred = [1, 0, 1, 1, 0, 0, 1, 1, 1, 0]
```

Utilizando o Scikit-learn, calcule a acurácia, precisão, recall e F1 do modelo.



```
from sklearn import metrics

y_true = [1, 0, 1, 0, 1, 0, 1, 1, 0, 0]
y_pred = [1, 0, 1, 1, 0, 0, 1, 1, 1, 0]

# Cálculo da acurácia
acuracia = metrics.accuracy_score(y_true, y_pred)
print("Acurácia:", acuracia)

# Cálculo da precisão
precisao = metrics.precision_score(y_true, y_pred)
print("Precisão:", precisao)

# Cálculo do recall
recall = metrics.recall_score(y_true, y_pred)
print("Recall:", recall)

# Cálculo do F1-score
f1_score = metrics.f1_score(y_true, y_pred)
print("F1-score:", f1_score)
```

Acurácia: 0.7
Precisão: 0.6666666666666666
Recall: 0.8
F1-score: 0.7272727272727272

R: Isso significa que o modelo acertou corretamente 70% das previsões, obteve precisão de 66%, um recall de 80% e um F1-score de 72%.