



Métricas de avaliação III: F1-Score

☰ Ciclo	Ciclo 06: Algoritmos baseado em árvores
# Aula	49
🕒 Created	@March 9, 2023 10:12 AM
☑ Done	☑
☑ Ready	☑

Objetivo da Aula:

- ☐ Revendo a Precision & Recall
- ☐ Precision/Recall TradeOff
- ☐ A métrica de F1-Score
- ☐ F1-Score na prática
- ☐ Resumo
- ☐ Próxima aula

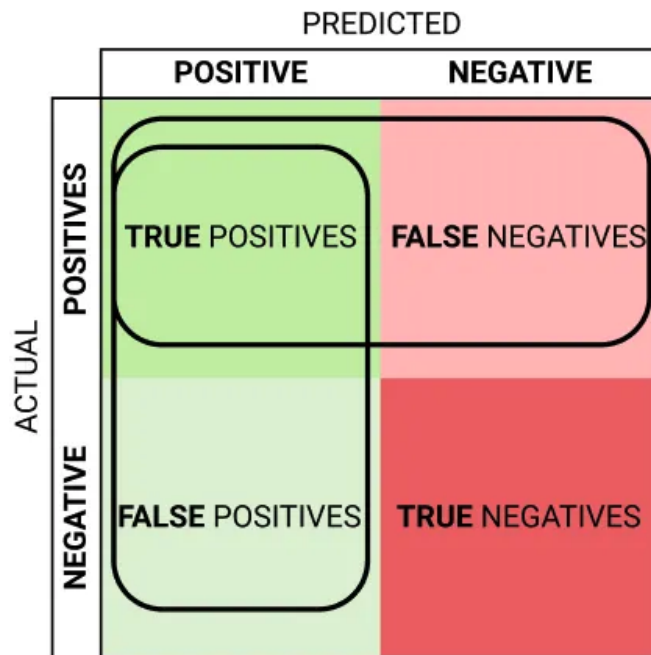
Conteúdo:

▼ 1. Revendo a Precision & Recall

Vamos imaginar que você reservou o sábado inteiro para ir ao pesqueiro. Abaixo estão os resultados da sua pescaria:

	Pescados	Total de peixes	Precision	Recall
Manhã	70	100	44%	70%
Tarde	30	30	19%	100%

▼ 1.1 Matriz de confusão



▼ Precisão

$$Precisão = \frac{TP}{TP + FP}$$

▼ Recall

$$Recall = \frac{TP}{TP + FN}$$

▼ 2. Precision/Recall TradeOff

Em alguns problemas de negócio, a métrica mais importante é a precisão, enquanto outros a métricas mais importante é recall.

Por exemplo, a empresa onde você trabalha desenvolve uma plataforma de hospedagem de vídeos de produtores de conteúdo com foco em crianças. Sua principal responsabilidade como Cientista de Dados da empresa, é detectar os vídeos seguros para crianças assistirem.

Nesse cenário, seria preferível um classificador que rejeitasse muitos bons vídeos classificando-os como “inapropriados”, mas conseguisse classificar todos os vídeos realmente inapropriados para menores como “inapropriados” (alta precisão).

Por outro lado, imagine que você trabalhe em uma empresa de vigilância e sua principal responsabilidade como Cientista de Dados é treinar um classificador para detectar ladrões em imagens de vigilância: provavelmente está tudo bem se seu classificador tiver apenas 30% de precisão, desde que tenha 99% de recall.

Nesse caso, os seguranças receberiam alguns alertas falsos, mas quase todos os ladrões serão pegos.

Em resumo, você não pode ter as duas métricas altas ao mesmo tempo, aumentar a precisão reduz o recall e vice-versa

Portanto, como escolher o valor para as métricas de Precisão e para o Recall? E se houvesse um meio termo?

▼ 3. A métrica de F1-Score

▼ 3.1 A média harmônica

A média harmônica é um tipo de medida estatística usada para calcular a média de grandezas inversamente proporcionais. Ela é calculada dividindo-se o número de elementos pela soma dos inversos desses elementos

▼ 3.1.1 Fórmula da média harmônica

$$Média Harmônica = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

Onde:

- n é o número de elementos
- x1, x2 ... são os valores dos elementos

▼ 3.2 F1-Score

A medida F1-score é a média harmônica da precisão e recall, uma vez que as duas medidas são inversamente proporcionais.

▼ 3.2.1 Fórmula da F1-score

$$F1 - score = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = 2 \times \frac{precision \times recall}{precision + recall}$$

▼ 3.3 Características do F1-Score

1. **Equilíbrio entre precisão e recall:** O F1-Score é uma medida que combina tanto a precisão quanto o recall de um modelo de classificação binária, particularmente útil em casos em que tanto a precisão

quanto o recall são importantes.

2. **Interpretação intuitiva:** O F1-Score é uma métrica fácil de interpretar, pois fornece um valor único que indica o desempenho geral do modelo. Além disso, o valor do F1-Score é compreensível para usuários não técnicos, o que pode ser útil em casos em que a avaliação do modelo precisa ser comunicada a pessoas que não têm conhecimento técnico.

▼ 3.4 Como interpretar o F1-Score

O intervalo do F1-Score varia de 0 a 1, onde 0 indica um desempenho muito ruim do modelo e 1 indica um desempenho excelente do modelo. O valor 0 é obtido quando tanto a precisão quanto o recall são zero, enquanto o valor 1 é obtido quando a precisão e o recall são iguais a 1.

Um valor alto de F1-Score indica que o modelo tem um bom equilíbrio entre a precisão e o recall, o que significa que ele está acertando muitos exemplos positivos e negativos. Por outro lado, um valor baixo de F1-Score indica que o modelo está tendo dificuldade em classificar corretamente as instâncias positivas ou negativas, ou que está tendo um desempenho desequilibrado em relação à precisão e ao recall.

▼ 4. F1-Score na prática

```
import numpy as np
from matplotlib import pyplot as plt

from sklearn.datasets import make_classification
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import precision_recall_curve, f1_score, auc, roc_curve, roc_auc_score, accuracy_score

# generate 2 class dataset
X, y = make_classification(n_samples=1000, n_classes=2, random_state=1)

# split into train/test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=2)

# fit a model
#model = RandomForestClassifier(n_estimators=100, max_depth=2, random_state=0)
model = LogisticRegression(solver='lbfgs')
model.fit(X_train, y_train)
yhat_test = model.predict( X_test )

# Performance
f1_score = f1_score( y_test, yhat_test )
print( "F1 Score: {}".format( f1_score ) )
```

▼ 5. Resumo

1. F1-Score é uma média harmônica usada para representar medidas que são inversamente proporcionais.
2. F1-Score avalia o equilíbrio entre precision e recall de um algoritmo de classificação.

▼ 6. Próxima aula

Exercícios

