



Métricas de avaliação I

# Aula	13
<input checked="" type="checkbox"/> Ready	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Finished	<input checked="" type="checkbox"/>
≡ Ciclos	Ciclo 02: Fundamentos

Objetivo da Aula:

- ☐ Métricas de avaliação de desempenho
- ☐ Quadro comparativo das métricas

Conteúdo:

▼ 1. Métricas de avaliação de desempenho

Avaliar o desempenho de modelos de séries temporais é essencial para garantir que as previsões sejam precisas e úteis para a tomada de decisão.

Como as séries temporais lidam com dados sequenciais e dependentes do tempo, as métricas de erro devem capturar a diferença entre os valores previstos e os valores reais, considerando tanto a magnitude do erro quanto seu impacto na interpretação dos resultados.

A seguir, estão as principais métricas utilizadas na avaliação de modelos de previsão de séries temporais:

▼ 1.1 Erro Médio Absoluto (Mean Absolute Error)

O **Erro Médio Absoluto (MAE)** mede a média das diferenças absolutas entre os valores previstos e os valores reais.

Ele é útil porque expressa o erro em unidades da variável analisada, tornando a interpretação intuitiva.

O MAE não diferencia erros negativos e positivos, pois considera apenas o valor absoluto do erro, evitando que erros opostos se cancelem.

Quanto menor o MAE, melhor o desempenho do modelo.

▼ 1.1.1. Fórmula do MAE

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

Onde:

- n é o número total de observações
- y_i é o valor real da observação i
- \hat{y}_i é o valor previsto para a observação i

▼ 1.2 Erro Quadrático Médio (Mean Square Erro)

O **Erro Quadrático Médio (MSE)** é uma métrica que calcula a média dos erros ao quadrado. A principal característica do MSE é penalizar erros maiores, pois os valores são elevados ao quadrado antes de serem somados.

Isso significa que grandes erros impactam significativamente essa métrica. Embora seja amplamente utilizado, seu valor não é facilmente interpretável na escala original dos dados, pois está em unidades ao quadrado.

▼ 1.2.1. Fórmula do MSE

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

Onde:

- n é o número total de observações
- y_i é o valor real da observação i
- \hat{y}_i é o valor previsto para a observação i

▼ 1.3 Raiz do Erro Quadrático Médio (RMSE)

A **Raiz do Erro Quadrático Médio (RMSE)** é a raiz quadrada do MSE e busca corrigir a escala dos erros ao trazê-los de volta para a unidade original dos dados. O RMSE também penaliza erros maiores, mas sua interpretação é mais intuitiva do que a do MSE.

Modelos com RMSE mais baixos apresentam melhor desempenho preditivo.

▼ 1.3.1. Fórmula do RMSE

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

Onde:

- n é o número total de observações
- y_i é o valor real da observação i
- \hat{y}_i é o valor previsto para a observação i

▼ 1.4 Erro Percentual Absoluto Médio (MAPE)

O **Erro Percentual Absoluto Médio (MAPE)** expressa o erro médio como uma porcentagem do valor real.

Ele é amplamente utilizado para avaliar previsões, pois fornece uma métrica relativa que facilita a comparação entre diferentes séries.

No entanto, o MAPE pode ser **problemático quando há valores reais próximos de zero**, pois a divisão por valores muito pequenos pode distorcer os resultados.

▼ 1.4.1. Fórmula do MAE

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Onde:

- n é o número total de observações

- y_i é o valor real da observação i
- \hat{y}_i é o valor previsto para a observação i

▼ 1.5 Critério de Informação de Akaike (AIC)

O **Critério de Informação de Akaike (AIC)** é uma métrica usada para comparar diferentes modelos, penalizando modelos mais complexos.

Ele leva em consideração tanto a qualidade do ajuste quanto o número de parâmetros do modelo, incentivando a escolha de modelos mais simples e eficientes.

Um valor de AIC menor indica um modelo melhor.

▼ 1.6 Critério de Informação Bayesiano (BIC)

Semelhante ao AIC, o **Critério de Informação Bayesiano (BIC)** também penaliza a complexidade do modelo, mas o faz de forma mais rígida.

Ele é especialmente útil quando se deseja evitar modelos excessivamente complexos que possam estar ajustando ruído em vez de padrões reais.

Assim como o AIC, um valor menor de BIC indica um modelo superior.

▼ 2. Quadro comparativo das métricas

▼ 1. MAE (Erro Médio Absoluto)

- **Descrição:** Mede o erro médio absoluto entre as previsões e os valores reais.
- **Interpretação:** Quanto menor, melhor; indica o erro médio na mesma unidade da variável.
- **Principal Característica:** Mede erro absoluto sem penalizar valores altos.
- **Pontos Fracos:** Não diferencia erros pequenos de grandes.
- **Pontos Fortes:** Simples e intuitivo, útil para entender o erro médio.
- **Ocasões de Uso:** Quando se deseja um erro interpretável na mesma unidade da variável.

▼ 2. MSE (Erro Médio Quadrático)

- **Descrição:** Calcula a média dos erros elevados ao quadrado.
- **Interpretação:** Quanto menor, melhor; valores altos indicam erros mais expressivos.
- **Principal Característica:** Penaliza erros maiores devido ao uso do quadrado.
- **Pontos Fracos:** Valores altos podem ser inflacionados por erros grandes.
- **Pontos Fortes:** Dá destaque a grandes erros, útil para identificar padrões ruins.
- **Ocasões de Uso:** Quando grandes erros devem ser penalizados fortemente.

▼ 3. RMSE (Raiz do Erro Médio Quadrático)

- **Descrição:** Raiz quadrada do MSE, mantendo a unidade original dos dados.
- **Interpretação:** Quanto menor, melhor; mede o erro médio na mesma unidade da variável.
- **Principal Característica:** Corrige a escala do MSE, tornando-o mais interpretável.
- **Pontos Fracos:** Ainda penaliza erros grandes mais do que pequenos.
- **Pontos Fortes:** Mais interpretável que o MSE e mantém a penalização de grandes erros.
- **Ocasões de Uso:** Quando se deseja um erro interpretável com penalização de grandes erros.

▼ 4. MAPE (Erro Percentual Médio Absoluto)

- **Descrição:** Mede o erro percentual médio absoluto em relação ao valor real.
- **Interpretação:** Quanto menor, melhor; expressa o erro como percentual do valor real.

- **Principal Característica:** Permite comparação entre séries de diferentes escalas.
- **Pontos Fracos:** Pode ser distorcido quando há valores reais próximos de zero.
- **Pontos Fortes:** Útil para comparar modelos em diferentes escalas e unidades.
- **Ocasões de Uso:** Quando é necessário comparar erros entre diferentes escalas.

▼ 5. **Bias (Viés do Modelo)**

- **Descrição:** Mede a média da diferença entre previsões e valores reais.
- **Interpretação:** Valores próximos de zero indicam que o modelo não superestima nem subestima.
- **Principal Característica:** Indica se o modelo tem tendência de erro sistemático.
- **Pontos Fracos:** Não quantifica a magnitude total do erro, apenas sua tendência.
- **Pontos Fortes:** Indica viés sistemático no modelo, ajudando a calibrá-lo.
- **Ocasões de Uso:** Para entender se o modelo tem tendência de super ou subestimar.

6. **AIC (Akaike Information Criterion - Critério de Informação de Akaike)**

- **Descrição:** Critério para avaliar modelos estatísticos baseado na qualidade do ajuste e complexidade.
- **Interpretação:** Menores valores indicam melhor modelo, penalizando a complexidade.
- **Principal Característica:** Equilibra qualidade do ajuste e simplicidade do modelo.
- **Pontos Fracos:** Depende do tamanho da amostra e não indica precisão absoluta.

- **Pontos Fortes:** Ajuda a comparar modelos balanceando ajuste e complexidade.
- **Ocasões de Uso:** Para comparar modelos estatísticos com diferentes parâmetros.

7. BIC (Bayesian Information Criterion - Critério de Informação Bayesiano)

- **Descrição:** Semelhante ao AIC, mas com uma penalização maior para modelos mais complexos.
- **Interpretação:** Menores valores indicam melhor modelo, penalizando fortemente a complexidade.
- **Principal Característica:** Favorece modelos mais simples e parcimoniosos.
- **Pontos Fracos:** Penaliza modelos mais complexos mais fortemente que o AIC.
- **Pontos Fortes:** Útil para escolher modelos estatísticos eficientes e sem overfitting.
- **Ocasões de Uso:** Quando se busca um modelo parcimonioso sem overfitting.