Métricas de avaliação II - Alg. Estatísticos

# Aula	14
Ready	[x]
Finished	[x]
Ciclos	Ciclo 02: Fundamentos

Objetivo da Aula:

- [ ] Métricas de avaliação de desempenho
- [ ] Quando usar cada tipo de métrica
- [ ] O que são algoritmos estatísticos?
- [ ] O que é verosimilhança?

Conteúdo:

1. Métricas de avaliação de desempenho

As métricas MAE, MSE, RMSE e MAPE têm como foco principal medir o **erro das previsões** do modelo. Elas avaliam a precisão das previsões ao comparar os valores reais da série temporal com os valores previstos pelo modelo. A principal preocupação dessas métricas é minimizar a diferença entre os valores observados e os valores previstos.

Diferentemente das métricas convencionais, **AIC** e **BIC** não medem diretamente a precisão das previsões, mas sim a **qualidade do ajuste do modelo estatístico** enquanto penalizam a complexidade do modelo.

1.5 Critério de Informação de Akaike (AIC)

AIC (Akaike Information Criterion) busca equilibrar o ajuste do modelo com o número de parâmetros usados. Ele penaliza modelos mais complexos, mas permite modelos um pouco mais sofisticados se melhorarem o ajuste dos dados. Sua fórmula é:

$$AIC = -2\ln(L) + 2k$$

Onde:

- L é a verossimilhança do modelo (o quão bem ele se ajusta aos dados)
- k é o número de parâmetros do modelo.

1.6 Critério de Informação Bayesiano (BIC)

BIC (Bayesian Information Criterion) também penaliza a complexidade do modelo, mas de forma mais rígida do que o AIC. Isso significa que o **BIC favorece modelos mais simples** do que o AIC. Sua fórmula é:

$$BIC = -2\ln(L) + k\ln(n)$$

2. Quando Usar Cada Tipo de Métrica?

- Use **MAE, MSE, RMSE e MAPE** quando o foco for avaliar a **precisão das previsões** do modelo em relação aos dados reais.
- Use **AIC e BIC** quando for necessário **comparar diferentes modelos estatísticos** para encontrar o melhor equilíbrio entre precisão e simplicidade, evitando modelos excessivamente complexos.

3. O que são algoritmos estatísticos?

Um **modelo estatístico** é uma representação matemática que descreve uma relação entre variáveis, baseada em princípios estatísticos e probabilísticos.

Todo modelo estatístico pode ser descrito por três componentes principais:

1. **Variáveis:**
    - Uma **variável dependente (ou resposta)**: o que queremos prever ou entender.
    - Uma ou mais **variáveis independentes (ou explicativas)**: fatores que influenciam a variável dependente.
2. **Relação Matemática entre as Variáveis**
    - Pode ser **determinística**, quando há uma relação fixa entre entrada e saída (como uma equação exata).
    - Pode ser **probabilística**, quando inclui incerteza e variabilidade (como um modelo de regressão, onde há erro associado às previsões).
3. **Suposições Estatísticas**
    - Os modelos estatísticos frequentemente fazem **suposições sobre os dados**, como normalidade, independência, estacionariedade, entre outras.

- Essas suposições ajudam a validar a aplicação do modelo e garantem que os resultados sejam confiáveis.

3.1. Qual a diferença entre modelos Estatísticos e de Machine Learning

- **Modelos Estatísticos**: Buscam entender relações entre variáveis e testar hipóteses. Eles são interpretáveis e baseados em suposições teóricas sobre os dados.
- **Modelos de Machine Learning**: Priorizam a capacidade preditiva, muitas vezes sem uma interpretação direta das relações entre as variáveis (por exemplo, redes neurais, florestas aleatórias)

Embora os modelos de Machine Learning possam ser baseados em princípios estatísticos, eles geralmente fazem menos suposições sobre os dados e se concentram mais na previsão do que na explicação dos fenômenos.

3.1.1. Quadro Resumo

Característica	Modelos Estatísticos	Modelos de Machine Learning
Objetivo	Explicação e previsão	Principalmente previsão
Interpretação	Fácil e direta (coeficientes explicam relações)	Muitas vezes difícil (modelos mais complexos)
Suposições sobre os dados	Muitas (linearidade, normalidade, independência dos erros)	Poucas ou nenhuma (aprendem padrões automaticamente)
Exemplos	Regressão Linear, ARIMA, Modelos Bayesianos	Árvores de Decisão, SVM, Redes Neurais
Flexibilidade	Limitada a relações matemáticas conhecidas	Muito flexível, capaz de capturar padrões complexos

4. O que é Verossimilhança?

A **verossimilhança** é um conceito estatístico que mede a **probabilidade dos dados observados terem sido gerados por um modelo específico**. Em termos simples, é uma função que indica **o quão bem um modelo estatístico se ajusta aos dados observados**.

Se tivermos um conjunto de dados X={x1,x2,...,xn} e um modelo parametrizado por θ, a **função de verossimilhança** é dada por:

$$L(\theta|X) = P(X|\theta)$$

Isso significa que **dado um conjunto de parâmetros θ\thetaθ, qual é a probabilidade de termos observado os dados X?**

4.1. Explicação simples da verossimilhança

Imagine que você levou uma criança no supermercado para fazer compras. Na gôndola de biscoitos, você sabe que a criança sempre escolhe sua caixa de chocolate favorita 8 vezes em cada 10, mas às vezes ele pega uma diferente.

Quando você chega em casa, descobre que algumas migalhas de chocolate estão no assoalho do carro e decide descobrir de qual caixa elas vieram.

Para isso, você reune as informações que possui:

1. **Caixa Azul**: Tem **80%** de chance de ser escolhida pela criança.
2. **Caixa Vermelha**: Tem **20%** de chance de ser escolhida pela criança.

A criança comeu **10 biscoitos**, e você encontrou **7 migalhas de biscoitos de chocolate** no carro. Agora você quer descobrir **de qual caixa os biscoitos vieram**.

Sabemos que:

- Na **Caixa Azul**, **90%** dos biscoitos são de chocolate, segundo o rótulo
- Na **Caixa Vermelha**, **50%** dos biscoitos são de chocolate, segundo o rótulo.

Qual a probabilidade das 7 migalhas de biscoito serem da caixa azul?

$$P(7 \text{ biscoitos de chocolate} \mid \text{Caixa Azul}) = \binom{10}{7}(0.9)^7(0.1)^3 = 0.057395628$$

Multiplicamos isso pela probabilidade de escolher a Caixa Azul (0.8):

$$L(\text{Caixa Azul}) = P(7|\text{Caixa Azul}) \times P(\text{Caixa Azul}) = 0.0459165$$

E qual a probabilidade das 7 migalhas de biscoito serem da caixa vermelha?

$$P(7 \text{ biscoitos de chocolate} \mid \text{Caixa Vermelha}) = \binom{10}{7}(0.5)^7(0.5)^3 = 0.1171875$$

Multiplicamos isso pela probabilidade de escolher a Caixa Vermelha (0.2):

$$L(\text{Caixa Vermelha}) = P(7|\text{Caixa Vermelha}) \times P(\text{Caixa Vermelha}) = 0.0234375$$

Conclusão

A verossimilhança da Caixa Azul (4.5%) é maior que a da Caixa Vermelha (2.3%). Isso significa que, com base nos dados observados (7 biscoitos de chocolate em 10), é mais provável que os biscoitos tenham vindo da Caixa Azul!

▼ 4.2. Explicação da verossimilhança na escolha dos modelos

Imagine que você tem um **cofrinho mágico** que gera números todos os dias. Seu trabalho é tentar adivinhar quais números vão sair amanhã! Você tem **dois amigos mágicos**, cada um com um método diferente para prever o número do dia seguinte:

1. **O Mago do ARIMA** 🔮 → Ele usa os números passados para calcular padrões e prever o próximo número.
2. **O Mago do Chute Aleatório** 🎲 → Ele simplesmente escolhe um número qualquer sem olhar os números anteriores.

Hoje, o cofrinho gerou o número **42**.

- 🔮 **Mago do ARIMA** diz que o número será **40**.
- 🎲 **Mago do Chute Aleatório** escolhe um número qualquer e diz que será **10**

Agora, queremos saber **qual dos dois magos fez a melhor previsão**.

A verossimilhança nos diz **quão boa foi cada previsão**. Como o número do cofrinho pode variar um pouco, vamos supor que ele segue uma **distribuição normal** com um pequeno erro, e um **desvio padrão de 5**.

A fórmula da distribuição normal para calcular a verossimilhança é:

$$L = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

onde:

- x = 42 (o número real gerado pelo cofrinho)
- μ é a previsão do mago
- σ=5 (desvio padrão que representa a incerteza)

Agora, aplicamos a fórmula para cada mago:

Mago do ARIMA:

$$L = \frac{1}{\sqrt{2\pi(5)^2}} e^{-\frac{(42-40)^2}{2(5)^2}} = 0.0737$$

Mago do Chute Aleatório:

$$L = \frac{1}{\sqrt{2\pi(5)^2}} e^{-\frac{(42-10)^2}{2(5)^2}} = 0.0000000001017$$

Resultado:

A verossimilhança do **Mago do ARIMA é milhões de vezes maior**, o que significa que **ele fez uma previsão muito melhor do que o chute aleatório**! 🎉

- 🔮 **Mago do ARIMA** tem **verossimilhança de 0.0737**.
- 🎲 **Mago do Chute Aleatório** tem **verossimilhança de 0.0000000001**.