

Documentação: ARIMA e SARIMA na Previsão de Vendas

Introdução

Modelos de séries temporais, como ARIMA (AutoRegressive Integrated Moving Average) e SARIMA (Seasonal AutoRegressive Integrated Moving Average), são ferramentas poderosas para prever dados sequenciais, como vendas mensais, temperatura diária ou índices financeiros. Esses modelos são amplamente utilizados para modelar séries temporais que apresentam dependência temporal, e sua aplicação no projeto de previsão de vendas mostrou-se eficaz.

O objetivo do projeto foi analisar o histórico de vendas de um produto e realizar previsões futuras com base nesse histórico. Para isso, foram utilizados os modelos ARIMA e SARIMA. A seguir, explicamos em detalhes o funcionamento de cada modelo e como eles foram aplicados no projeto.

1. ARIMA (AutoRegressive Integrated Moving Average)

O **ARIMA** é um modelo de séries temporais que combina três componentes principais para modelar dados com dependências temporais:

- **AR (AutoRegressivo - AR)**: Representa a relação entre uma observação e um número de observações anteriores. Em outras palavras, a variável dependente é regida por uma combinação linear de seus próprios valores passados.
- **I (Integrated - I)**: Refere-se ao processo de diferenciação para tornar a série temporal estacionária. Séries temporais que não são estacionárias precisam ser transformadas até que elas se tornem estacionárias.
- **MA (Moving Average - MA)**: Representa a relação entre uma observação e o erro cometido nas previsões anteriores. Esse componente modela o erro como uma média ponderada de erros passados.

A notação do ARIMA é dada como **ARIMA(p, d, q)**, onde:

- **p** é o número de termos auto-regressivos,
- **d** é o número de diferenciações para tornar a série estacionária,
- **q** é o número de termos de média móvel.

Como Funciona o ARIMA?

1. **Estacionarização (Diferenciação)**: A primeira etapa do modelo ARIMA é garantir que a série temporal seja estacionária. Isso é feito por meio da diferenciação, subtraindo valores sucessivos da série. Se a série ainda não for estacionária após uma diferenciação, outras diferenciações podem ser realizadas.
2. **Modelagem AR (AutoRegressiva)**: Após garantir a estacionariedade da série, o modelo AR identifica a relação entre o valor da série e seus próprios valores passados. O parâmetro **p** especifica quantos valores passados da série devem ser usados para prever o valor futuro.

3. **Modelagem MA (Média Móvel):** O componente MA do modelo ARIMA modela o erro da previsão como uma média ponderada de erros passados. O parâmetro q define quantos termos de erro passado são usados.

Exemplo de ARIMA

Para aplicar o ARIMA no histórico de vendas de um produto, seguimos os seguintes passos:

1. **Tornar a série estacionária:** Se a série apresenta tendência (por exemplo, aumento ou diminuição de vendas ao longo do tempo), utilizamos diferenciação (subtração de valores sucessivos) para torná-la estacionária.
2. **Escolher os parâmetros p , d , q :** A partir da série estacionária, escolhemos os parâmetros ideais utilizando análise de autocorrelação (ACF e PACF) ou experimentação de valores.
3. **Ajustar o modelo ARIMA:** Com os parâmetros escolhidos, ajustamos o modelo ARIMA à série temporal de vendas.
4. **Avaliar o modelo:** Após a previsão, comparamos os resultados obtidos com os valores reais, utilizando métricas como o erro quadrático médio (RMSE) e a média absoluta dos erros (MAE).

2. SARIMA (Seasonal ARIMA)

O **SARIMA** é uma extensão do ARIMA que lida com séries temporais sazonais. Muitas séries temporais, como vendas de produtos ou temperatura mensal, têm padrões sazonais, ou seja, comportamentos cíclicos que se repetem em intervalos regulares. O SARIMA adiciona a capacidade de modelar esses padrões sazonais.

A principal diferença entre ARIMA e SARIMA é a adição de componentes sazonais. Enquanto o ARIMA trata apenas de dependências temporais simples, o SARIMA captura os efeitos sazonais, como ciclos anuais, mensais, etc.

O modelo SARIMA é representado como **SARIMA(p, d, q)(P, D, Q, m)**, onde:

- (p, d, q) são os parâmetros do modelo ARIMA convencional (AutoRegressivo, Integrado e Média Móvel).
- (P, D, Q) são os parâmetros sazonais:
 - P : Ordem da parte auto-regressiva sazonal.
 - D : Ordem de diferenciação sazonal.
 - Q : Ordem da parte de média móvel sazonal.
- m é o número de períodos na sazonalidade. Por exemplo, para dados mensais com sazonalidade anual, $m = 12$.

Como Funciona o SARIMA?

1. **Identificação da Sazonalidade:** A primeira etapa é identificar o período sazonal da série temporal, por exemplo, $m = 12$ para dados mensais com sazonalidade anual.
2. **Modelagem Sazonal:** O SARIMA adiciona os parâmetros sazonais P, D, Q para capturar padrões sazonais, além dos parâmetros tradicionais p, d, q do ARIMA.
3. **Diferenciação Sazonal:** O parâmetro D realiza uma diferenciação sazonal para tornar a série estacionária, considerando os padrões sazonais. Por exemplo, para dados mensais, a

diferenciação seria a subtração entre o valor de um mês e o mês correspondente no ano anterior.

4. **Modelagem AR e MA Sazonais:** O modelo SARIMA também captura dependências temporais e erros passados (AR e MA), mas leva em conta os padrões sazonais com os parâmetros **P** e **Q**.

Exemplo de SARIMA

Suponha que temos uma série temporal de vendas mensais de um produto com um ciclo sazonal anual (por exemplo, aumento de vendas no final do ano). Podemos usar o modelo **SARIMA(1, 1, 1)(1, 1, 1, 12)**, onde:

- **ARIMA(1, 1, 1)** modela a dependência temporal simples,
- **(1, 1, 1, 12)** modela a sazonalidade anual, considerando um ciclo de 12 meses.

3. Aplicação no Projeto: Previsão de Vendas

No contexto do projeto de previsão de vendas de um produto, a aplicação dos modelos ARIMA e SARIMA seguiu as etapas descritas a seguir:

1. **Análise do Histórico de Vendas:** Foi coletado o histórico de vendas mensais do produto, que foi analisado para verificar a presença de tendência e sazonalidade. Em caso de padrões sazonais claros (como aumento de vendas no final do ano), o modelo **SARIMA** foi utilizado. Caso contrário, o **ARIMA** foi suficiente para modelar os dados.
2. **Escolha do Modelo:**
 - Para dados sem sazonalidade, utilizamos o **ARIMA**.
 - Para dados com sazonalidade (padrões anuais de vendas, por exemplo), o **SARIMA** foi utilizado.
3. **Pré-processamento dos Dados:** A série temporal foi transformada para garantir sua estacionariedade, aplicando diferenciação quando necessário, e identificando possíveis componentes sazonais.
4. **Treinamento e Validação do Modelo:** O modelo foi treinado utilizando os dados históricos, ajustando os parâmetros com base em técnicas de análise de autocorrelação e validação cruzada. Foram utilizadas métricas como RMSE e MAE para avaliar o desempenho do modelo.
5. **Previsão de Vendas:** Após ajustar e validar o modelo, realizamos a previsão das vendas futuras do produto. As previsões fornecidas pelos modelos ARIMA e SARIMA foram analisadas para auxiliar na tomada de decisões sobre estoques e estratégias de marketing.

4. Conclusão

Os modelos ARIMA e SARIMA são ferramentas poderosas para previsão de séries temporais. O **ARIMA** é ideal para séries sem sazonalidade, enquanto o **SARIMA** é uma extensão que captura padrões sazonais, tornando-o mais adequado para dados com ciclos regulares. A aplicação desses modelos no projeto de previsão de vendas proporcionou insights valiosos e ajudou a planejar com maior precisão as vendas futuras do produto.

Esses modelos são amplamente utilizados em diversos setores, como finanças, economia e marketing, e podem fornecer previsões confiáveis quando ajustados corretamente. A escolha entre ARIMA e SARIMA depende da natureza dos dados: se apresentam ou não sazonalidade.