



Universidade Presbiteriana
Mackenzie

Inteligência Artificial

módulo 4

Introdução a
Previsão de Séries
Temporais e a
Aprendizagem por
Reforço

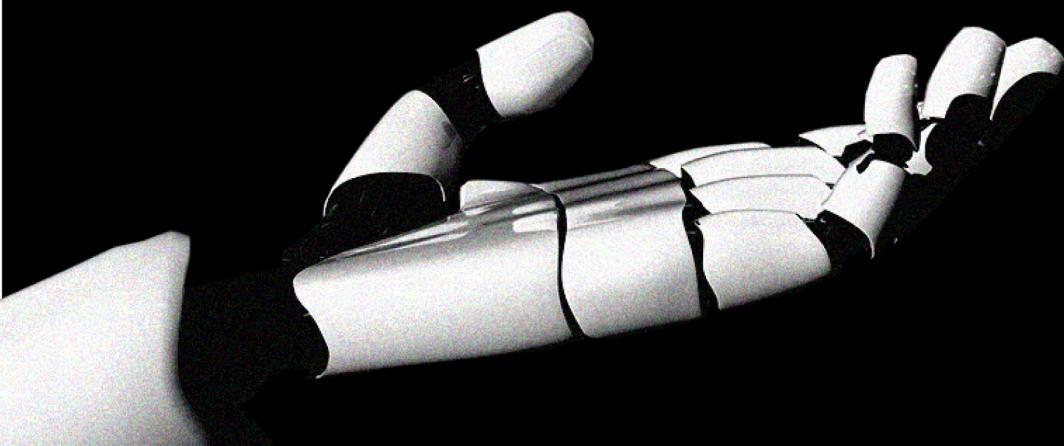


Universidade Presbiteriana
Mackenzie



Trilha 1

Introdução a Séries Temporais



Sumário

1	Apresentação do componente curricular	p. 4
2	Introdução ao estudo da trilha de aprendizagem	p. 6
3	Introdução a séries temporais e aplicações	p. 7
4	Principais técnicas de previsão	p. 14
5	Síntese	p. 19
6	Referências	p. 20

Apresentação do componente curricular

O componente curricular “Introdução à Previsão de Séries Temporais e à Aprendizagem por Reforço” foi elaborado com o objetivo de introduzir o tema de séries temporais e de aprendizagem por reforço para os estudantes de Inteligência Artificial, assim apresentando assuntos muito relevantes que podem ser utilizados em diversas áreas de atuação. Este componente foi dividido em oito trilhas de aprendizagem, das quais as quatro iniciais são referentes às séries temporais e as quatro finais ao aprendizado por reforço. Os tópicos abordados em cada trilha são os seguintes:

Trilha 1: “Introdução a séries temporais”

Conceitos gerais de séries temporais, sua aplicabilidade e as principais técnicas de previsão.

Trilha 2: “Tendência e sazonalidade”

Apresentação dos conceitos de tendência e sazonalidade vinculados às séries temporais.

Trilha 3: “Médias móveis”

Conceituação deste tópico muito importante quando se estuda uma série temporal, as médias móveis. Seu conceito e técnicas de uso.

Trilha 4: “Regressão para séries temporais e suavização”

Finalização dos estudos sobre séries temporais, apresentando os principais conceitos de regressão para séries temporais e de suavização.

Trilha 5: “Introdução à aprendizagem por reforço”

Principais conceitos de aprendizagem por reforço e os algoritmos mais relevantes.



Trilha 6: “Aplicações de aprendizagem por reforço”

O problema da Aprendizagem por Reforço, exemplos de aplicabilidade.

Trilha 7: “Programação dinâmica e Q-Learning”

Conceituação sobre a Programação Dinâmica e apresentação do Q-Learning.

Trilha 8: “Deep Learning para Aprendizagem por reforço”

Apresentação sobre como utilizar o Deep Learning ao se trabalhar com Aprendizagem por Reforço.

Introdução ao Estudo da Trilha de Aprendizagem

O objetivo desta trilha é discutir os principais conceitos e técnicas referentes ao estudo de séries temporais, vinculados à Inteligência Artificial, tendo como base os seguintes tópicos:

- introdução às Séries Temporais;
- exemplos de aplicações do uso de séries temporais;
- principais técnicas de previsão.

Introdução a séries temporais e aplicações

Quando vamos estudar séries temporais, podemos imaginar que são assuntos novos que surgiram com as tecnologias recentes, como aprendizagem de máquina ou inteligência artificial. Porém um dos primeiros trabalhos que mencionam as séries temporais foi apresentado nos anos 20 do século passado por Yule (1927).

Neste trabalho publicado, Yule estudou as manchas solares onde aplicou um modelo autorregressivo linear, no qual o valor previsto baseava-se nos valores anteriores.

Ao longo das décadas, muitos outros pesquisadores se aprofundaram sobre a temática das séries temporais. Dentre os diversos estudos, um que se destacou é o material apresentado por Box & Jenkins (1976), no qual reuniram as principais técnicas e metodologias relacionadas às séries temporais que tratavam de previsão e estimação de parâmetros. Box & Jenkins consideravam que a combinação dos valores precedentes e os impactos aleatórios anteriores com o impacto atual apresentava o valor previsto da série temporal.

Quase duas décadas depois da publicação de Box e Jenkins, Chaves (1991) apresentou a necessidade de se considerar os modelos não lineares, já que a maioria das séries reais apresentam uma grande tendência da não linearidade.

Dentre as técnicas mais comumente utilizadas para efetuar previsões, uma das mais aplicadas considera as observações ordenadas cronologicamente da variável que se deseja prever (CHATFIELD, 2000).



Quando falamos sobre estudos e análises de séries temporais, consideramos que existe uma relação constante entre os eventos que ocorreram no passado e que podem continuar a influenciar os dados futuros. Estes eventos não são aleatórios, assim, é possível realizar procedimentos estatísticos e aplicar modelos.

Uma definição formal da análise de séries temporais é a de possuímos um conjunto de informações sobre uma variável, obtido em períodos regulares e ordenados no tempo, e pretendermos prever os valores futuros desta série (MORETTIN & TOLOI, 1987).

Morettin e Toloi (1987, 2006) também introduziram uma forma de se analisar as séries temporais através da **decomposição, considerada clássica, da série em quatro componentes: ciclo, tendência, sazonalidade e a variação aleatória**. Nem sempre a decomposição de uma série temporal irá apresentar estes quatro componentes:

- **Ciclos:** os ciclos são as oscilações de subida e descida encontradas nas séries temporais; eles podem ser suaves e repetidos, acompanhando o componente de tendência. Os movimentos cíclicos normalmente ocorrem em intervalos irregulares de tempo. Na figura 1, vemos um exemplo de ciclo:

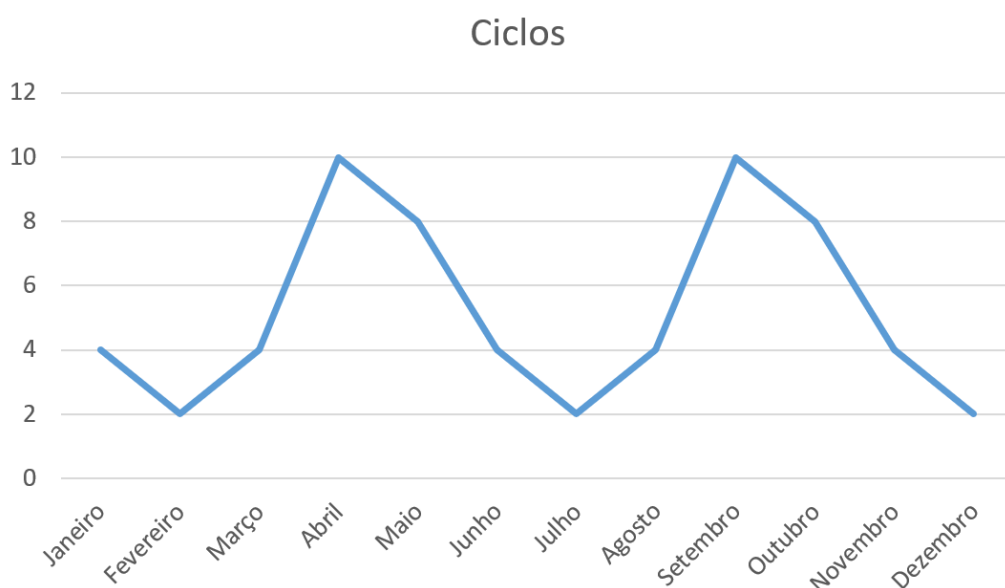
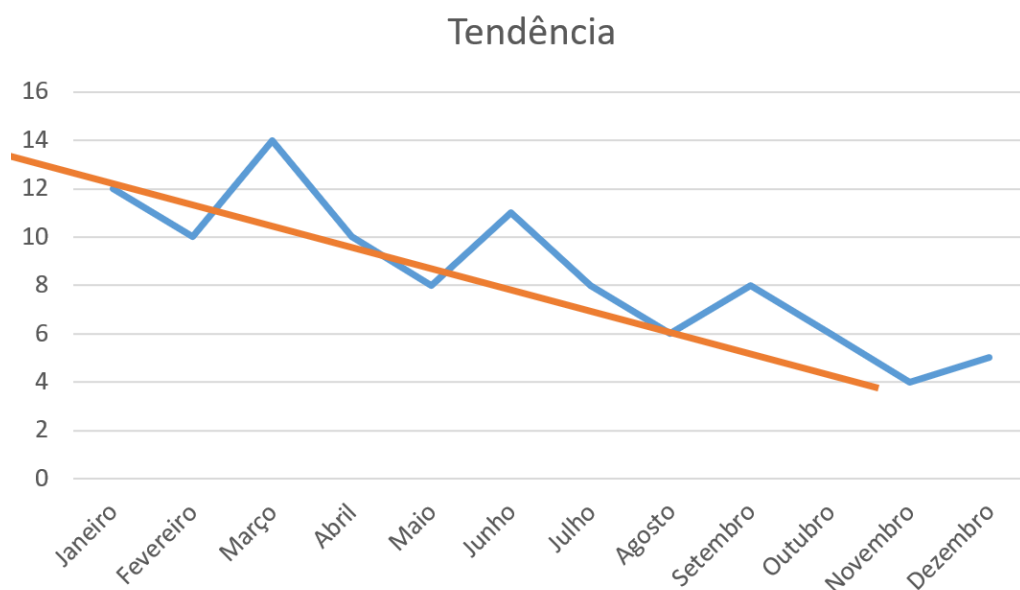


Figura 1 - Ciclos em uma série temporal.
Fonte: Elaborado pelo autor.

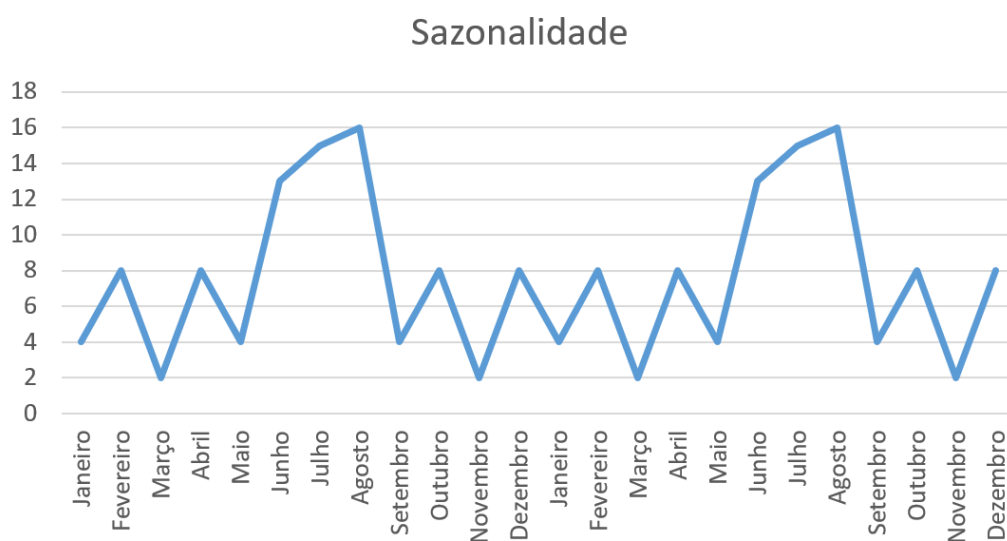
- **Tendência:** a tendência indica se a série temporal pode crescer, decrescer ou se manter estável ao longo do tempo, além de indicar a velocidade com que as variações ocorrem. Vemos um exemplo de tendência decrescente na figura 2:

Figura 2 - Tendência em uma série temporal. Fonte: Elaborado pelo autor.



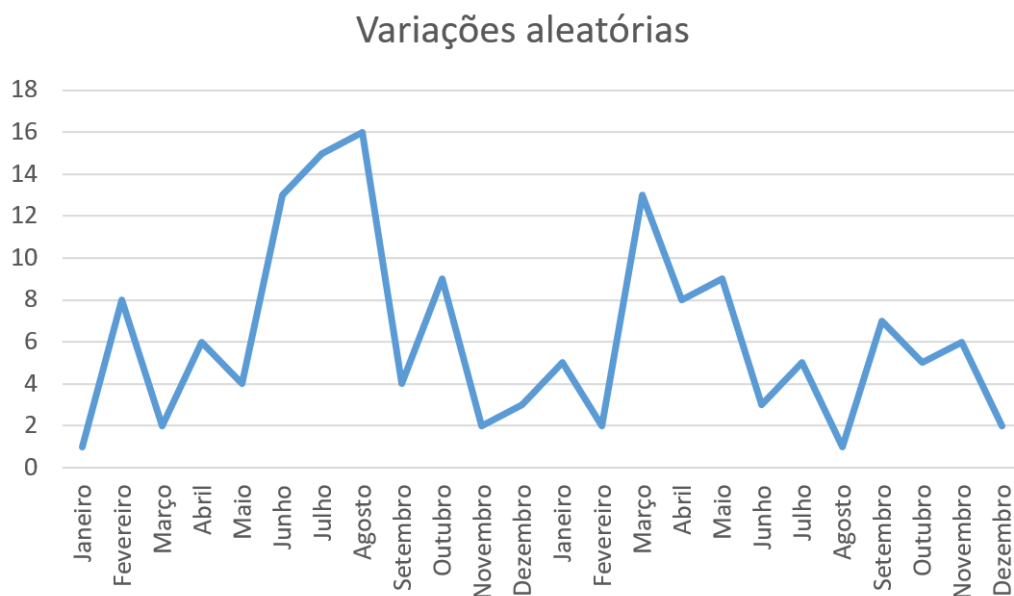
- **Sazonalidade:** quando uma série temporal apresenta oscilações de subida e de descida que ocorrem sempre em um determinado período, como ano, mês ou dia da semana, temos o componente da sazonalidade. A sazonalidade possui movimentos que ocorrem em intervalos regulares de tempo, assim sendo mais fácil de prevê-los. A figura 3 apresenta um exemplo de sazonalidade em uma série temporal:

Figura 3 - Sazonalidade em uma série temporal. Fonte: Elaborado pelo autor.



• **Variação aleatória:** são as variações que os demais componentes não conseguem representar. É quando a série temporal apresenta variações muito grandes de valores, com muitas flutuações que podem ocorrer devido a algum fator aleatório. Este componente também pode ser chamado de erro aleatório. A figura 4 ilustra um exemplo de variação aleatória.

Figura 4 - Variáveis aleatórias em uma série temporal.
Fonte: Elaborado pelo autor.



Uma representação da equação para se obter a série temporal é:

$$Y_t = f(T_t, C_t, S_t, \varepsilon_t)$$

Onde temos que Y_t é o valor da série temporal e:

- T_t é a componente Tendência no período t ,
- C_t é a componente Ciclo no período t ,
- S_t é a componente de Sazonalidade no período t ,
- ε_t é a componente de erro aleatório ou ruído branco no período t .

O componente Ciclo acaba representando movimentos com períodos longos, porém, como normalmente os períodos observados são pequenos em comparação com o tamanho do ciclo, então muitas vezes o que se observa, quando se analisa o efeito da Tendência, é uma parte do Ciclo.

Morettin e Toloi (2006) consideram que, como na maioria das vezes não podemos diferenciar o efeito do componente Ciclo do componente Tendência, então subtrai-se o Ciclo da composição da série temporal. Outro ponto importante é que nem sempre uma série temporal irá apresentar todos os componentes.

Assim uma série temporal pode ser representada da seguinte forma, considerando a decomposição clássica:

$$X_t = T_t + S_t + a_t, t = 1, 2, \dots, N$$

onde T_t é o componente de tendência, S_t é o componente cíclico ou sazonal e a_t é a componente variável aleatória ou ruído branco.

Aplicação das séries temporais

!

Quando se possui uma série temporal, a expectativa é de que exista uma causa vinculada ao tempo que influenciou os dados no passado e que pode influenciá-los no futuro. Identificar estas causas possibilita estimar com precisão as variáveis antes que elas ocorram. Com isto, a aplicação prática das análises de séries temporais é muito ampla.

Morettin e Toloi (2006) apresentam alguns dos principais objetivos para se estudar e analisar as séries temporais, estes objetivos são vistos a seguir:

- **Descrição:** a descrição é verificar o padrão de comportamento da série temporal utilizando, para isto, gráficos, verificando se tem os componentes de tendência, ciclos e sazonalidade além de outras ferramentas descritivas;
- **Modelagem:** a modelagem se baseia na elaboração de modelos a fim de se conseguir compreender o mecanismo gerador da série temporal. Com isto, é possível utilizar a variação de uma série temporal para analisar e justificar a variação em outra série. Para que

o modelo possa ser desenvolvido, é preciso definir os parâmetros e também a avaliação da qualidade estatística para o ajuste do modelo;

- **Previsão:** é quando se usam os valores passados para prever ou estimar os valores futuros de uma série temporal. As aplicações de previsão podem ser feitas considerando prazos variáveis, desde curtos até longos prazos;

- **Controle de processos:** Os parâmetros referentes à qualidade de um processo se encontram registrados em uma série temporal, assim, é possível monitorar estes valores e com isto detectar mudanças estruturais na série. Quando se notam alterações estruturais, uma das possibilidades é que o modelo atual utilizado para retratar o funcionamento do processo se encontra desatualizado.

Com estes objetivos, as aplicações das séries temporais não ficam restritos a poucas áreas, como econômica e financeira, mas também podem ser empregados em áreas como: Marketing, Produção, Gestão Empresarial e Recursos Humanos (MAKRIDAKIS & WHEELWRIGHT, 1989).

Segundo Box e colaboradores (1994), temos que as séries temporais podem ser aplicadas também na estimativa de valores futuros de uma série temporal, considerando os valores passados e os atuais, e na determinação de um modelo de entrada-saída (input-output) que possibilite realizar análises utilizando qualquer série temporal como entrada e verificando os efeitos com os resultados de saída.

Os estudos e os avanços propostos pelas análises das séries temporais são muito amplos. Desde que se tenham os dados, as possibilidades de aplicações são muito grandes. Temos estudos, por exemplo, referentes à área de epidemiologia (ANTUNES & CARDOSO, 2015) e até sobre a cobertura de vegetação em determinadas regiões (BORGES & SANO, 2014). Como as técnicas e metodologias de análises das séries

temporais se encontram bem elaboradas, muitos estudiosos têm aplicado os conhecimentos em diversos setores.

Principais técnicas de previsão

As técnicas de previsão variam consideravelmente, e são desenvolvidas com vários propósitos distintos. Cada técnica possui características próprias, grau de precisão e custo de utilização, os quais devem ser considerados na escolha de um método específico (PELLEGRINI & FOGLIATTO, 2001).

Conforme apresentado por Morettin e Toloí (2006), todos os métodos de previsão são apenas diferentes procedimentos para calcular a previsão de mínimos quadrados de um valor futuro, considerando para isto uma combinação linear de valores passados.

!

Com isto, a previsão de um valor futuro deve ser considerada como uma informação complementar para uma tomada de decisão. Ou seja, o valor futuro apresentado não deve ser considerado como o valor que realmente irá ocorrer no futuro, mas sim servir como base para auxiliar nas decisões a serem tomadas.

Quanto aos métodos de previsão, estes podem ser classificados em três tipos (CHAFTFIELD, 2000):

- **Previsões *Judgemental*:** este tipo de previsão não considera qualquer informação relevante, mas apenas o julgamento subjetivo, a intuição;
- **Métodos univariados:** é quando, para se obterem as previsões, são utilizados somente os valores passados de uma série temporal. Neste caso, pode-se utilizar uma função do tempo ou colocar uma tendência linear para auxiliar. Nesta classificação, temos os modelos de autorregressão (AR), de médias móveis (MA), ajustamento sazonal e alisamento exponencial e os modelos autorregressivos integrados e de médias móveis (ARIMA);



- **Métodos multivariados:** é quando se modelam simultaneamente duas ou mais séries temporais sem que exista qualquer obrigatoriedade de ter uma relação entre elas.

Para Morettin e Tolo (2006), os modelos de estudos das séries temporais podem ser classificados conforme os números de parâmetros:

- **Modelos paramétricos:** número de parâmetros finitos. Nestes modelos, a análise é feita no domínio do tempo e os principais modelos adotados são os modelos de regressão, os modelos autorregressivos (AR), os modelos de média móvel (MA), os modelos autorregressivos e de médias móveis (ARMA) e os modelos autorregressivos integrados e de médias móveis (ARIMA);
- **Modelos não paramétricos:** número infinito de parâmetros. Estes modelos utilizam análises no domínio de frequência, assim eliminam o problema da correlação serial. Consideram os modelos de função de autovariância (ou autocorrelação) e sua transformação de Fourier.

Neste componente, abordaremos o modelo clássico com as análises através da decomposição da série temporal em componentes de tendência e sazonalidade, veremos a aplicação da média móvel e, por fim, as técnicas de regressão.

Porém, conforme mencionado anteriormente, existem outros modelos para estudos e análises das séries temporais, são eles:

Modelos Autorregressivos (AR)

!

O modelo autorregressivo (AR, do inglês *Auto-Regressive*) é um modelo estocástico que utiliza uma combinação linear entre os valores do passado da série temporal e um valor randômico, assim apresentando o resultado esperado para o valor futuro (BOX et al., 2015).



Morettin e Toloi (2006) apresentam que o modelo AR deve possuir os parâmetros do modelo relacionados com um ruído branco. O ruído branco tem como características possuir média zero, variância constante e ter autocorrelação igual à zero (BUENO, 2008).

A autocorrelação pode ser de primeira ordem ou de ordem q , e corresponde à correlação que existe entre uma observação i e a observação i anterior. Sendo a observação i anterior $(i-1)$ quando for de primeira ordem e $(i-q)$ quando for de ordem q .

Assim a variável futura a ser prevista através do modelo AR é a resultante dos valores que são combinações lineares dos p valores passados mais um ruído a_t .

Modelos de Média Móvel (MA)

!

O modelo de médias móveis (do inglês *Moving Average*) realiza uma média ponderada dos valores dos ruídos observados nos períodos passados. Este modelo considera que a série temporal a ser modelada é formada por uma combinação linear de q sinais de ruídos, os quais são independentes entre si.

O modelo MA considera que a série é finita e temos uma constante, os parâmetros do modelo e um ruído branco. Para compreender a diferença entre o modelo AR e o modelo MA, é preciso entender o conceito de autocorrelação.

A **autocorrelação** é quando temos uma associação entre os valores de uma mesma variável. Normalmente os valores podem ser ordenados no tempo, utilizando os dados das séries temporais. Existem dois tipos de funções para se trabalhar com a autocorrelação: a função de autocorrelação e a de autocorrelação parcial.

A função de autocorrelação é utilizada para avaliar como a quantidade de períodos da série vinculados a uma observação precedente (defasagem) pode influenciar no valor atual, já

a função de autocorrelação parcial possui como objetivo filtrar correlações de outros intervalos de tempo, mantendo somente a correlação pura entre duas observações.

Um diferencial entre o modelo de médias móveis com relação ao modelo autorregressivo é que o modelo MA utiliza a função de autocorrelação, enquanto o modelo AR utiliza a função autocorrelação parcial.

Modelos Autorregressivos e de Médias Móveis (ARMA)

Quando utilizamos o modelo de médias móveis (MA) e o modelo autorregressivo (AR) separadamente, precisamos adequar os parâmetros para cada modelo e os resultados obtidos por ambos nem sempre são os melhores possíveis. Os modelos autorregressivos e de médias móveis (ARMA – do inglês *Auto-Regressive Moving Average*) consideram a integração entre os modelos AR e MA, atuando na mesma série temporal, diminuindo a quantidade de parâmetros e com isto apresentando uma melhora no ajuste.

O modelo ARMA, além de possuir menos parâmetros para se trabalhar, apresenta resultados mais precisos. Este modelo que inclui os termos autorregressivos e de médias móveis foi proposto como uma alternativa para simplificar um modelo com grande número de parâmetros (MORETTIN & TOLOI, 2006).

Modelos autorregressivos integrados e de médias móveis (ARIMA)

!

Os modelos autorregressivos integrados e de médias móveis (ARIMA, do inglês *Auto-Regressive Integrated Moving Average*) são modelos para análise de séries temporais, sendo uma generalização de modelos ARMA, em que o termo “integrado” é devido a um processo de diferenciação da série temporal original transformando-a em estacionária, com isto, o ARIMA possibilita trabalhar com séries temporais não estacionárias.

Uma série temporal estacionária tem como característica as propriedades estatísticas constantes ao longo do tempo, como exemplo, podemos citar a média e a autocorrelação. Já em contrapartida, uma série não estacionária possui variações ao longo do tempo de suas propriedades estatísticas.

Síntese

Neste *e-book*, foram apresentados os principais conceitos sobre séries temporais, percorrendo a sua história e passando pelos seus componentes principais; tendência, ciclo, sazonalidade e variável aleatória. Vimos as possibilidades de uso da análise de séries temporais e que as possibilidades são várias, em todas as áreas de negócio. Por fim, apresentamos as principais técnicas para trabalhar com as séries temporais e suas características.

Esta trilha possui um fórum, no qual vocês poderão compartilhar seus conhecimentos e ideias. Continuem consumindo os outros materiais da trilha, como o audioblog e o vídeo explicativo. O conjunto de todos estes materiais possibilitarão uma compreensão melhor do que é uma série temporal e suas aplicações.



Referências

AGGARWAL, C. C. **Neural Networks and Deep Learning: A textbook**. Berlin: Springer, 2018.

ANTUNES, J. L. F.; CARDOSO, M. R. A. Uso da análise de séries temporais em estudos epidemiológicos. **Epidemiologia e Serviços de Saúde**, v. 24, p. 565-576, 2015.

BISGAARD, S.; KULAHCI, M. **Time Series Analysis and Forecasting by Example**. Wiley, 2011.

BORGES, E. F.; SANO, E. E. Séries temporais de EVI do MODIS para o mapeamento de uso e cobertura vegetal do oeste da Bahia. **Boletim de Ciências Geodésicas**, Curitiba, v. 20, n.3, p. 526-547, setembro 2014.

BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis, Forecasting and Control**. Holden Day, 1976.

BOX, G. E. *et al.* **Time series analysis: forecasting and control**. John Wiley & Sons, 2015.

BUENO, R. L. S. **Econometria de Séries Temporais**. Editora Cengage Learning, 2008

CHAFTFIELD, C. **Time Series Forecasting**. Chapman & Hall/CRC, 2000.

CHAVES, A. N. **Bootstrap em Séries Temporais**. Rio de Janeiro. 208f. Tese (Doutorado em Engenharia Elétrica), Universidade Federal do Rio de Janeiro, 1991.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn and TensorFlow**. Sebastopol: O'Reilly, 2017.



GRAESSER, L.; KENG, W. L. **Foundations of Deep Reinforcement Learning: Theory and Practice in Python**. Boston: Addison-Wesley, 2019.

MAKRIDAKIS, S.; WHEELWRIGHT, S. **Forecasting Methods for Management**. 5^a ed. New York: John Wiley & Sons, 1989.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting**. 2^a ed. Wiley, 2015.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2^a ed. São Paulo: Edgard Blücher/ ABE- Projeto Fisher, 2006.

MORETTIN, P. A.; TOLOI, C. M. C. **Previsão de séries temporais**. 2^a ed. São Paulo: Atual, 1987.

PELLEGRINI, F. R.; FOGLIATTO, F. S. Passos para implantação de sistemas de previsão de demanda: técnicas e estudo de caso. **Produção**, São Paulo, v. 11, n. 1, p. 43-64, junho 2001.

SANTOS, G. Q. V.; MARQUES JUNIOR, J. A.; BERNARDO, Y. N. S. Previsão de Demanda: Revisão bibliográfica e análise acadêmica atual. **XXXV Encontro Nacional de Engenharia de produção (ENEGEP)**, 2015, Fortaleza. Anais ENEGEP, 2015.

SUTTON, R. S.; BARTO, A. G. **Reinforcement Learning**. 2^a ed. Cambridge: The MIT Press, 2018.

YULE, G. U. *On a method of investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers*. **Philos. Trans. R. Soci.**, v. 226, n. 636-646, 1927.



