

# Séries Temporelles

24/02/2022



# Cronograma

- Modelos Híbridos para séries temporais
- Como estruturar um processo de forecast
- Lidando com previsões múltiplas
- ARMA models e Prophet
- Avaliação de Modelos
- Aplicação: forecast simplificado usando Darts



# O que são séries temporais?

Uma série temporal pode ser definida como um **conjunto de observações ordenadas no tempo**, como, por exemplo:

- Os valores mensais de temperatura registrados na cidade de São Paulo;
- O salário mensal de uma pessoa durante um ano;
- O preço de ações em um determinado período de tempo.



# Modelos Híbridos (finalização da última aula)

Existem algoritmos de predição que apresentam um bom desempenho em estimar a tendência de uma série temporal. Existem modelos que se saem bem identificando sazonalidades e ciclos.

Portanto, vê-se a utilidade em criar modelos híbridos que contam com ambos os algoritmos: um para identificar (e extrapolar) a tendência, e outro para identificar sazonalidade



# Modelos Híbridos

Transformamos a série temporal em

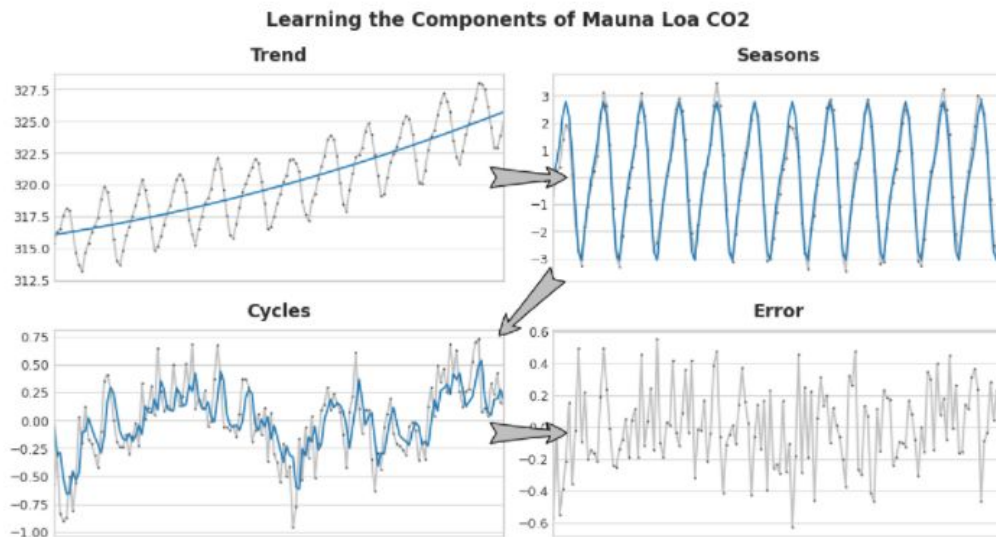
- $\text{Série} = \text{Tendência} + \text{Sazonalidade} + \text{Erro}$

Passo a Passo

- 1- Fitar um modelo para aprender a tendência
- 2 - Subtrair a tendência estimada da série
- 3 - Com os valores resultantes, fitar um modelo para a sazonalidade
- 4 - O que sobra disso seria nosso Erro



# Modelos Híbridos (esquema)



*Learning the components of Mauna Loa CO2 step by step. Subtract the fitted curve (blue) from its series to get the series in the next step.*



# Processo de Forecasting

Em posse de muitas informações e características sobre séries temporais, é necessário entender quais as etapas para a resolução de um problema de forecasting.

Para isso, precisamos entender os conceitos de **forecast origin**, **forecast horizon** e **lead time**

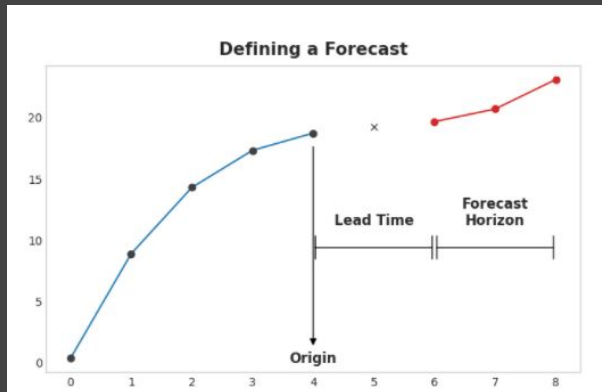


# Processo de Forecasting

**Forecast Origin:** o instante no qual você realiza o forecast, tudo até esse ponto será usado como treino.

**Forecast Horizon:** horizonte de tempo o qual será predito.

**Lead Time:** Tempo entre a origem e o início do forecast horizon





# Processo de Forecasting

Após a definição do problema, partimos para o pré-processamento e criação de features, como vimos semana passada.

Date	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
2016-01-04	0.0	0.0	0.0	0.0	0.0	0.0
2016-01-05	1.0	0.0	0.0	0.0	0.0	0.0
2016-01-06	0.0	1.0	0.0	0.0	0.0	0.0
2016-01-07	0.0	0.0	1.0	0.0	0.0	0.0
2016-01-08	0.0	0.0	0.0	1.0	0.0	0.0
2016-01-09	0.0	0.0	0.0	0.0	1.0	0.0
2016-01-10	0.0	0.0	0.0	0.0	0.0	1.0
2016-01-11	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...

	Hardcover	Lag_1
Date		
2000-04-01	139	NaN
2000-04-02	128	139.0
2000-04-03	172	128.0
2000-04-04	139	172.0
2000-04-05	191	139.0

	Hardcover	Time
Date		
2000-04-01	139	0
2000-04-02	128	1
2000-04-03	172	2
2000-04-04	139	3
2000-04-05	191	4



# Obs: Previsões múltiplas

Para a realizar uma tarefa de forecast que envolva a previsão de mais de um valor (forecast horizon  $> 1$ ) é necessário adotar algumas estratégias, uma vez que nem todos os modelos realizam previsões múltiplas.

Nesse sentido, vamos analisar algumas dessas estratégias:

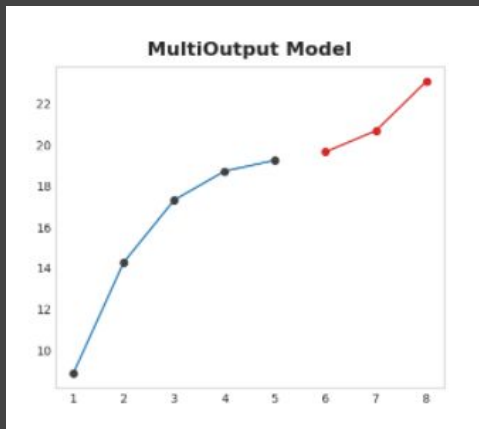
- Multi Output models
- Direct Strategy
- Recursive Strategy



# Obs: Previsões múltiplas

## Multi Output models

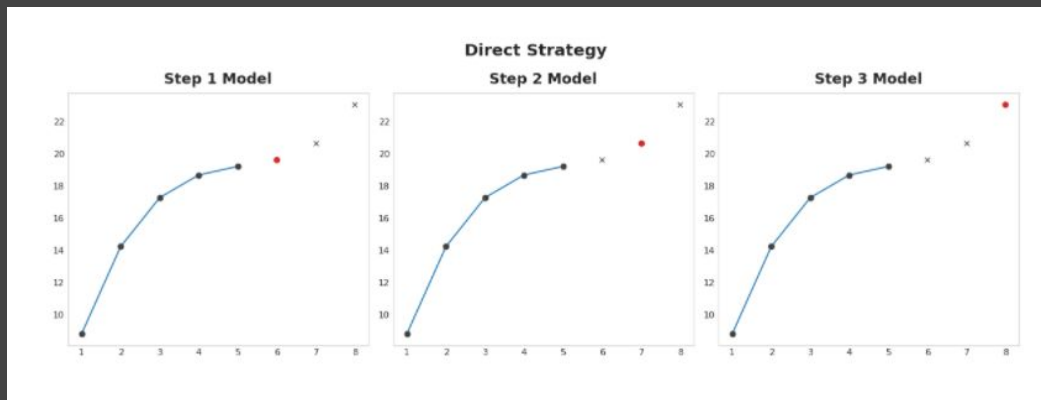
- Usar algoritmos e modelos que, por definição, aceitam múltiplos outputs (Regressão Linear, Redes Neurais).



# Obs: Previsões múltiplas

## Direct Strategy

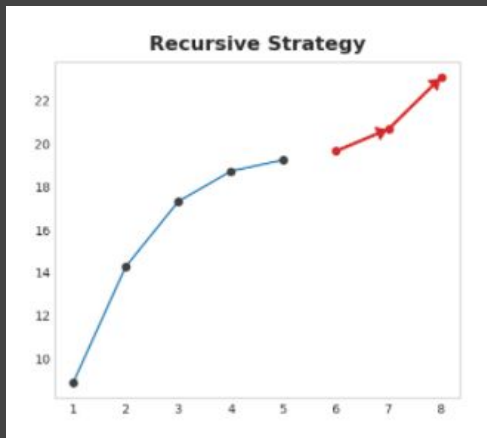
- Treinar um modelo separado para cada ponto a ser predito (muito “caro” computacionalmente para um horizonte de forecast grande).



# Obs: Previsões múltiplas

## Recursive Strategy

- Realizar uma predição e utilizar o valor predito como feature para gerar a predição seguinte.



# Modelos ARMA (ARMA, ARIMA, SARIMA)

Os modelos ARMA (AutoRegressive Moving Average) são geralmente utilizados no mundo de time series forecast e tem sua base teórica pautada na transformação da série temporal estudada em termos de dois polinômios:

- Polinômio para auto regressão: regressão usando os próprios valores da série de forma defasada
- Polinômio para a média móvel: tenta modelar os termos de erro

Muitas pessoas se referem a ARMA como  $ARMA(p,q)$ , em que  $p$  é a ordem do polinômio auto regressivo e  $q$  a ordem do polinômio da MA.

Problema: ARMA básico só funciona com dados estacionários



# Modelos ARMA: ARIMA e SARIMA

- ARIMA (AutoRegressive Integrated Moving Average): aparece para fornecer uma maneira de lidar com dados não estacionários. Adicionando um polinômio diferenciável para representar a estacionariedade.
- SARIMA (Seasonal ARIMA): adiciona sazonalidade com mais polinômios.

Ideia principal: modelar a sua série temporal como uma soma de polinômios



# Facebook's Prophet

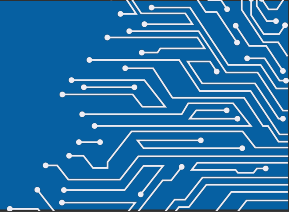
Modelo de predição de time series criado pelo Facebook

*“Modelo aditivo onde tendências não lineares são ajustadas junto a sazonalidades anuais, semanais e diárias, além de efeitos de feriados. Funciona melhor com séries temporais com efeitos sazonais fortes e muitas temporadas de dados históricos” Sr. Facebook*





# Métricas de Avaliação



Por fim, após a criação dos modelos, é necessário avaliá-los, para isso, utiliza-se algumas métricas na tentativa de quantificar se os resultados produzidos são suficientes para o projeto realizado.

OBS: Não existe a **melhor métrica** na avaliação de um modelo. Cada valor representa e enfatiza um aspecto diferente dos seus resultados, a tarefa principal consiste em selecionar as métricas mais adequadas para avaliar e diferenciar os modelos produzidos.



# Métricas de Avaliação

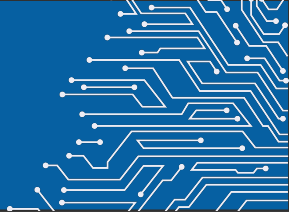
## Erro Absoluto Médio (MAE)

- Média das diferenças absolutas
- Fácil de entender, fácil de explicar
- Não pune outliers
- Bom para casos em que o projeto é focado em identificar tendências, não os valores precisos na série.

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



# Métricas de Avaliação



## Erro Quadrático Médio (MSE)

- Média do erro das previsões ao quadrado
- Predições muito distantes do real aumentam o valor da medida muito facilmente
- Bom para punir outliers

## Raiz do Erro Quadrático Médio (MSE)

- Tira a raiz da métrica de cima com o objetivo de fornecer um valor com a mesma unidade da série temporal



# Métricas de Avaliação

## Erro Percentual Absoluto Médio (MAPE)

- Média da divisão da diferença entre predito ( $\hat{y}$ ) e real pelo valor real ( $y$ )
- Exprime uma porcentagem
- Fácil interpretação
- Muito usada para séries temporais

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



# Aplicação Usando Darts

Biblioteca do python especializada em forecasts de séries temporais

