

Análise de Algoritmos de Aprendizagem com Base em Séries Temporais

Escola Superior de Tecnologia e Gestão - P.Porto Licenciatura em Engenharia Informática Projeto Final 2019/20



Estrutura







Contextualização

Objetivos

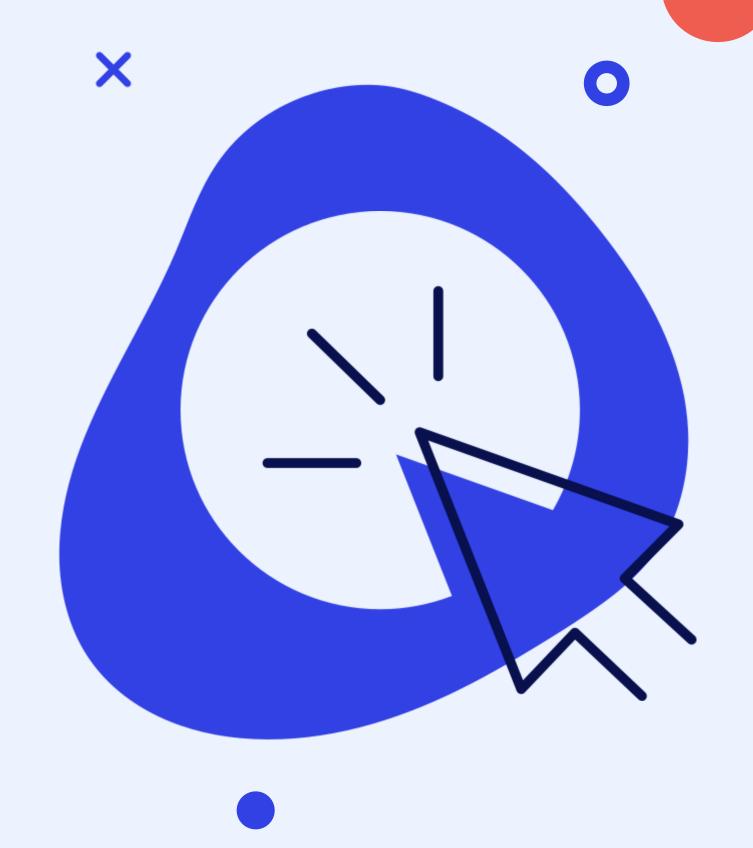
Estudar **séries temporais**

Estudar alguns **algoritmos de aprendizagem**

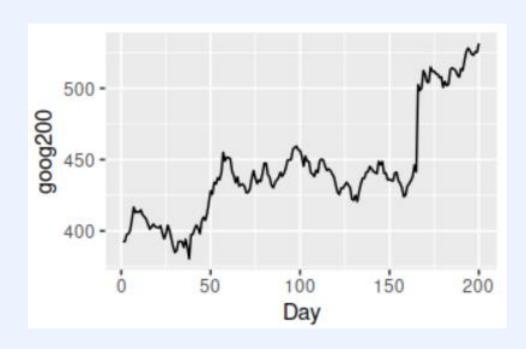
Desenvolver uma **plataforma de testes** de modelos

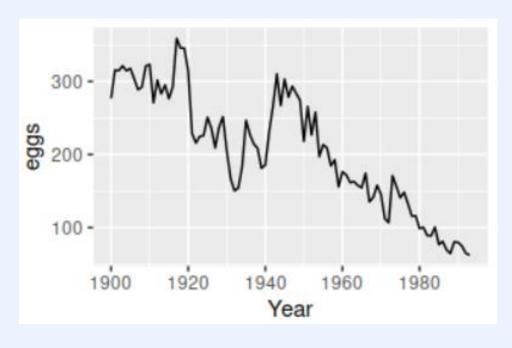
Realizar alguns casos de estudo

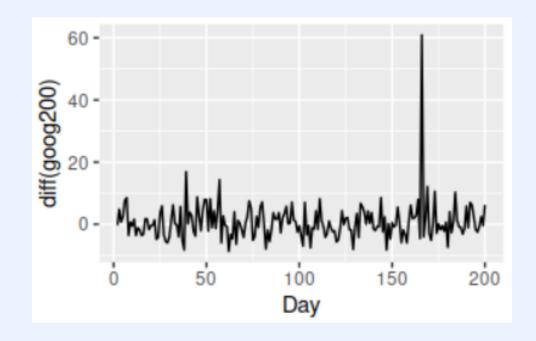
Adaptar para o ambiente do Google Colab

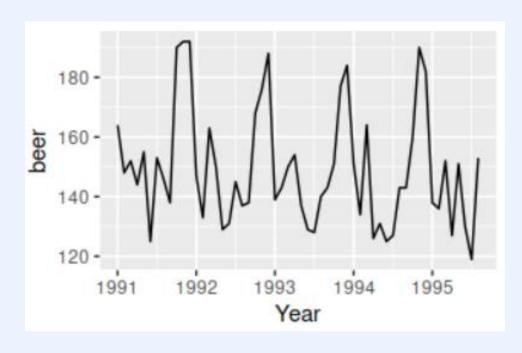


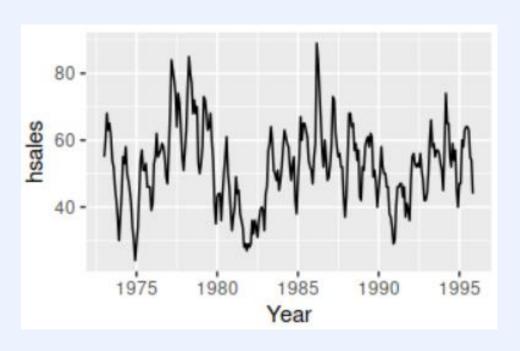
Séries Temporais

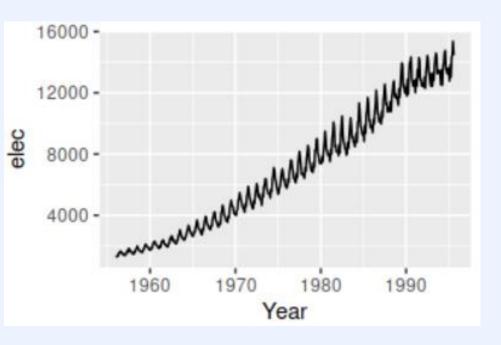














Séries Temporais

"date"	"census"	
01/01/2016	4092	
02/01/2016	3807	
03/01/2016	3208	
04/01/2016	1057	
05/01/2016	3302	
06/01/2016	3761	
07/01/2016	2964	

"date"	"wifi status"	"temp"	"precipitation"	"census"
01/01/2016	up	-6	0	4092
02/01/2016	up	2	0	3807
03/01/2016	up	2	0	3208
04/01/2016	up	-4	0	1057
05/01/2016	up	2	0	3302
06/01/2016	up	2	0	3761
07/01/2016	up	4	0	2964

Modelos de Aprendizagem

(ARIMA, ARIMAX, SARIMA e SARIMAX)

ARIMA

AR: Autoregressão

I: Integrado

MA: Média Móvel

ARIMA(p, d, q)

p: Ordem de Atraso

d: Grau de Diferenciação

q: Ordem da Média Móvel



Modelos de Aprendizagem

(ARIMA, ARIMAX, SARIMA e SARIMAX)

ARIMAX

X: Variáveis Exógenas

SARIMA

S: Sazonal

SARIMA(p, d, q)(P, D, Q, S)

P: Ordem de Atraso Sazonal

D: Grau de Diferenciação Sazonal

Q: Ordem da Média Móvel Sazonal

S: Temporada



Google Colab

Promover e incentivar o estudo de Machine Learning

Plataforma com planos gratuitos

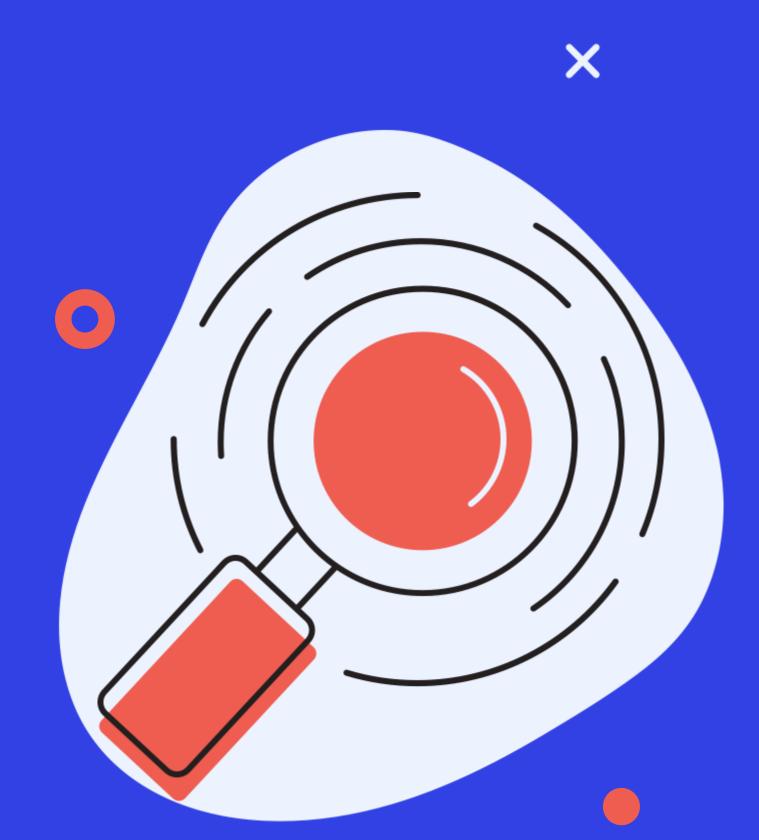
Correr os algoritmos em **máquinas potentes**

Integração com o Google Drive









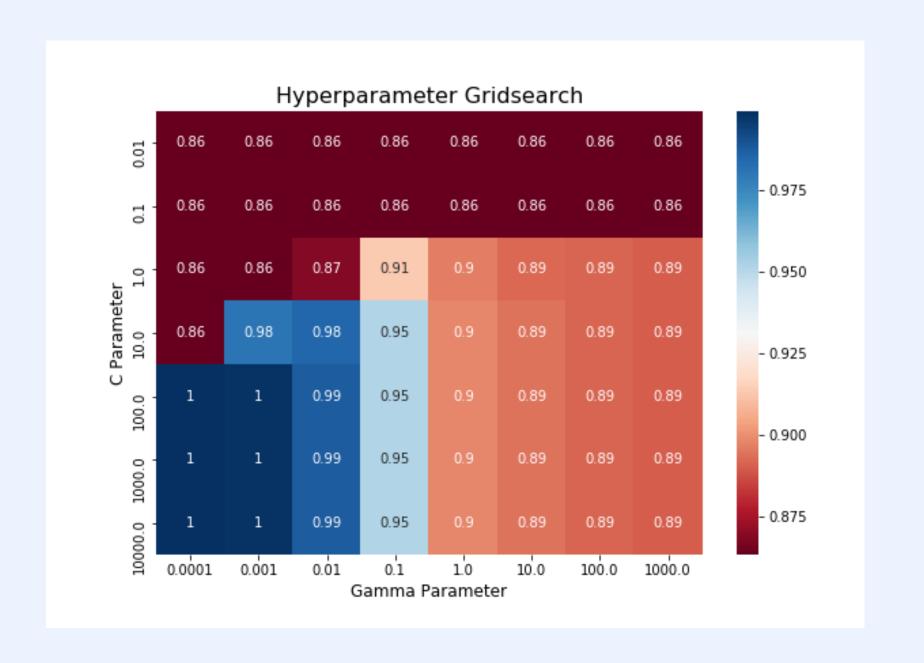
Casos de Estudo

Grid Search

Pode ser realizada em **qualquer tipo de modelo**.

Constrói um modelo para **cada combinação** de parâmetros possível.

Pode ser **extremamente cara** e levar bastante tempo.





Grid Search

ARIMA(p, d, q) com:

- p entre 1 e 5;
- d entre 0 e 2;
- q entre 0 e 2.

Escolher 4 melhores configurações do ARIMA.

SARIMA(p, d, q)(P, D, Q, S) com:

- Pentre 0 e 5;
- Dentre 0 e 2;
- Q entre 0 e 2;
- S sendo 24.

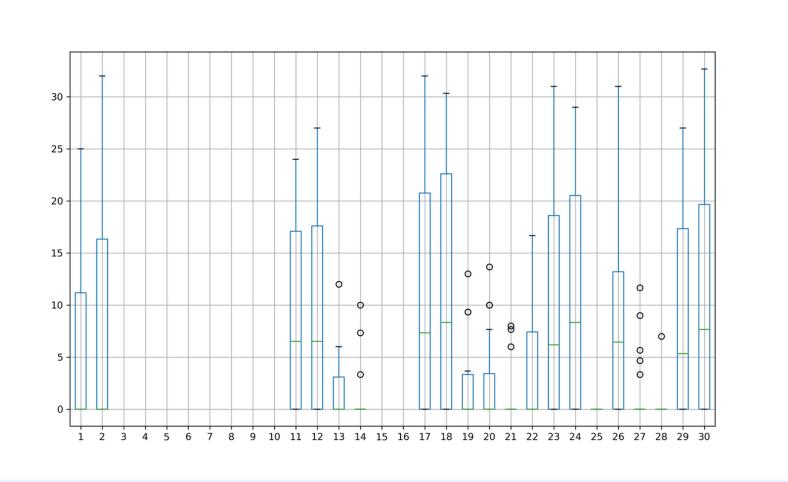
Escolher 2 melhores configurações do SARIMA.

Escolher as variáveis exógenas que melhores resultados obtiveram para os modelos ARIMAX e SARIMAX.



Velocidades Instantâneas

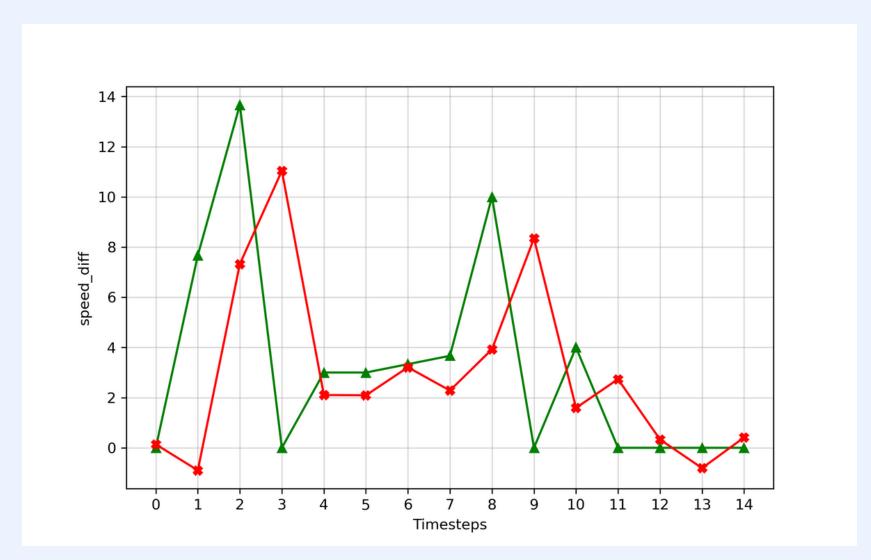
"timestep"	"temperature"	"precipitation"	"speed_diff"
2019-04-01 07:00:00	12.0	0.0	11.66666666666668
2019-04-01 08:00:00	15.0	0.0	0.0
2019-04-01 09:00:00	16.0	0.0	11.0
2019-04-01 10:00:00	17.0	0.0	19.33333333333333
2019-04-01 11:00:00	18.0	0.0	21.333333333333333
2019-04-01 12:00:00	18.0	0.0	17.66666666666668
2019-04-01 13:00:00	16.0	0.0	25.0



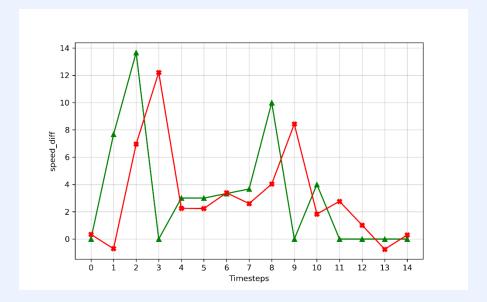


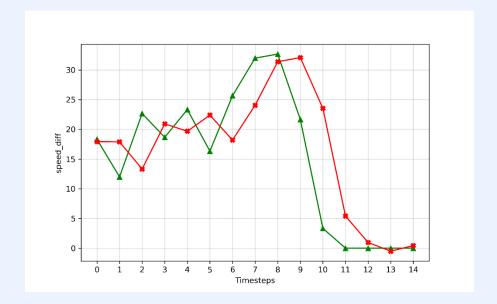
Velocidades Instantâneas

Questão: Qual a melhor e a pior hora para circular na estrada de carro?



ARIMAX(1, 2, 3), 1ª divisão do dataset, com 15 previsões

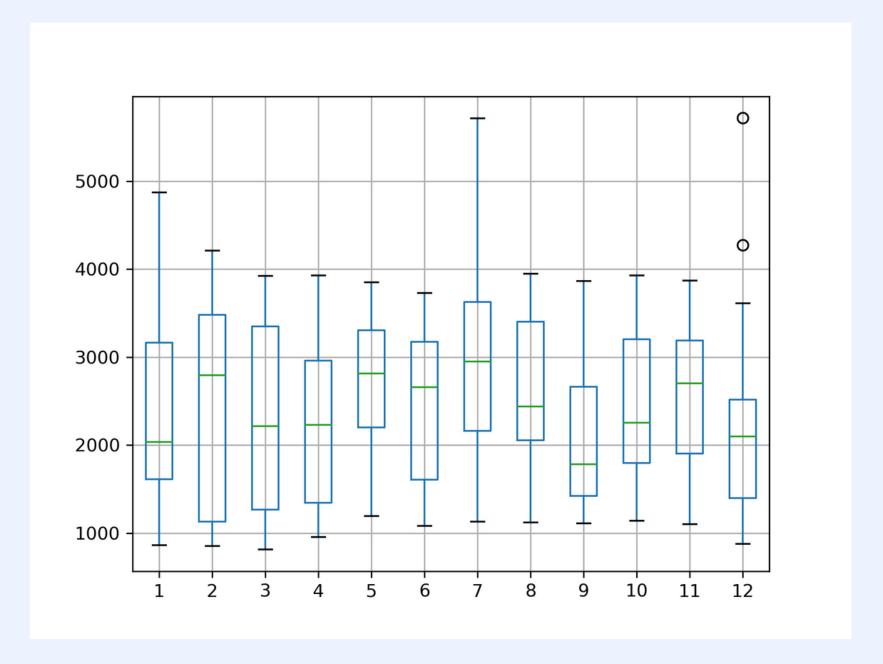






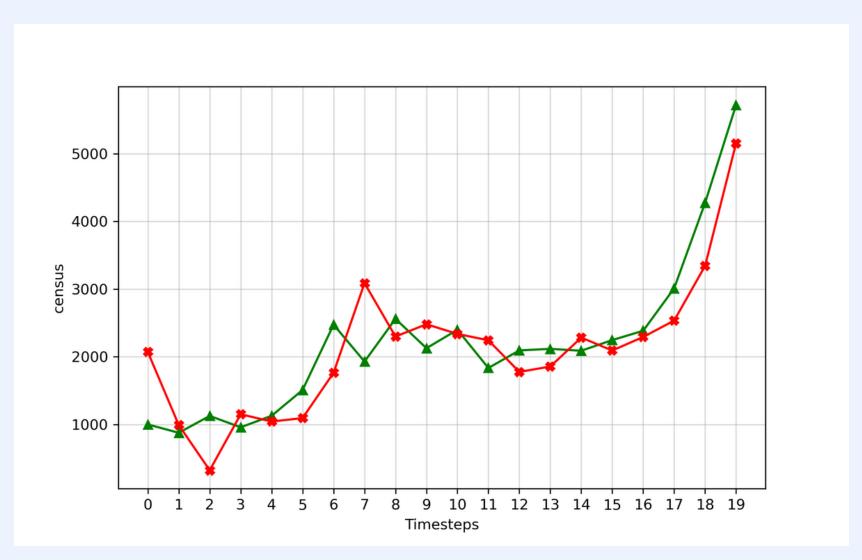
Quantidade Populacional

"date"	"wifi status"	"temp"	"precipitation"	"census"
01/01/2016	up	-6	0	4092
02/01/2016	up	2	0	3807
03/01/2016	ир	2	0	3208
04/01/2016	up	-4	0	1057
05/01/2016	ир	2	0	3302
06/01/2016	ир	2	0	3761
07/01/2016	up	4	0	2964

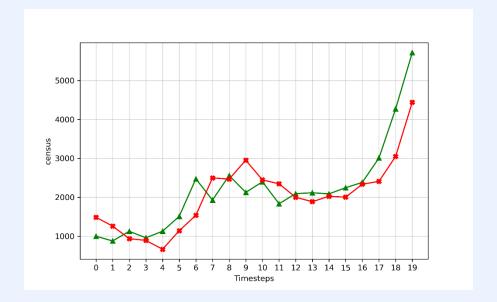


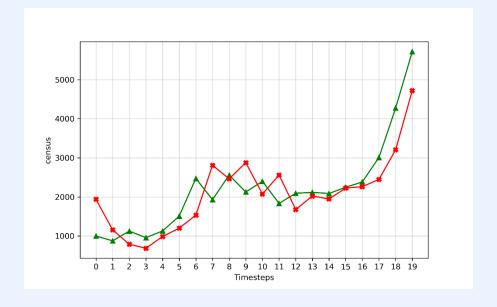
Quantidade Populacional

Questão: Qual o melhor e o pior dia do mês para ir ao quiosque?

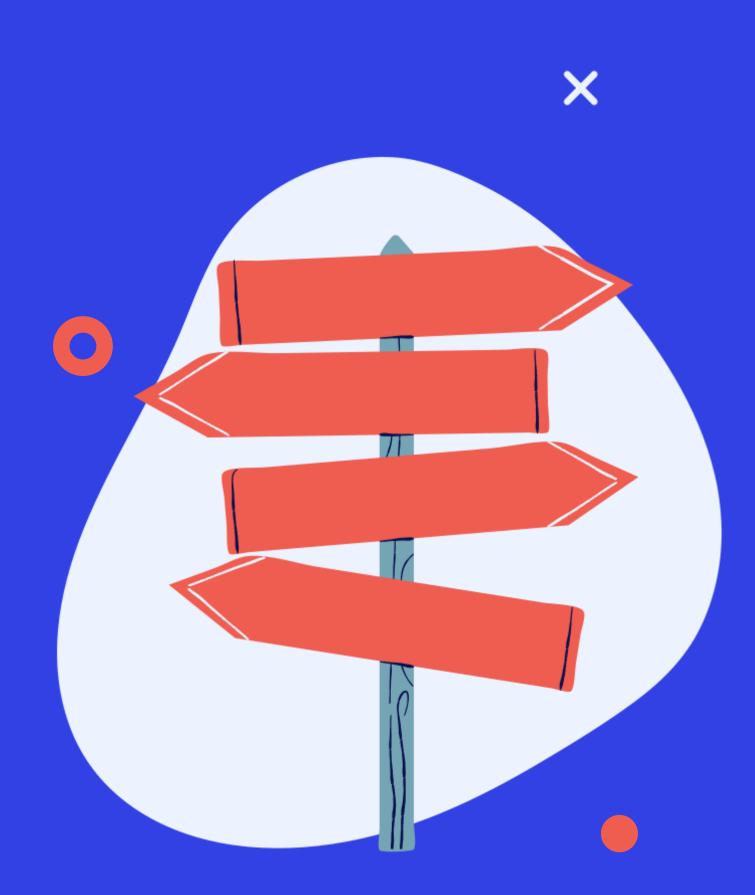


ARIMA(1, 2, 0), sem validação cruzada, com 20 previsões









Trabalho Futuro



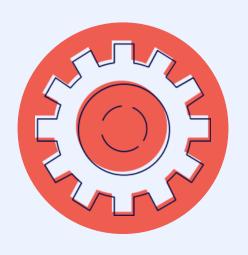
Trabalho Futuro



Migração para nova biblioteca



Implementação do ambiente num servidor



Servir API através de servidor



Obrigado!



Repositório

https://github.com/LuisMarques99/ProjetoFinal

