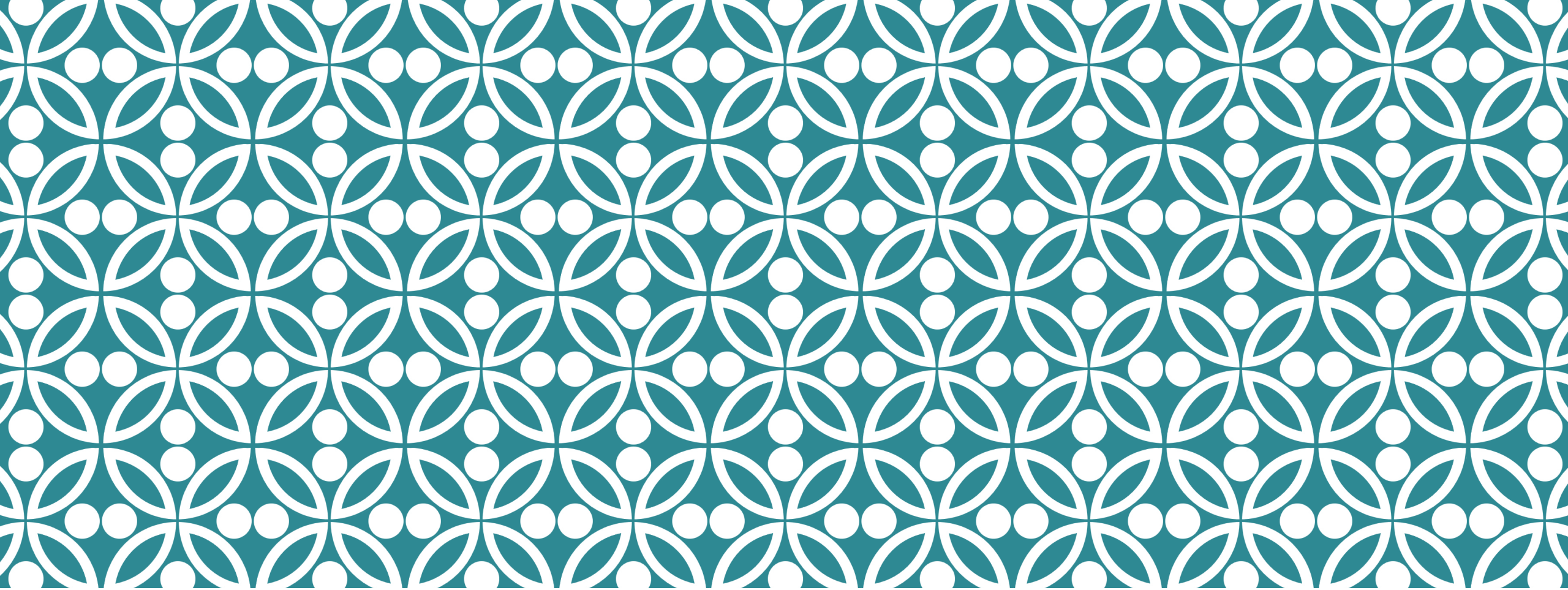
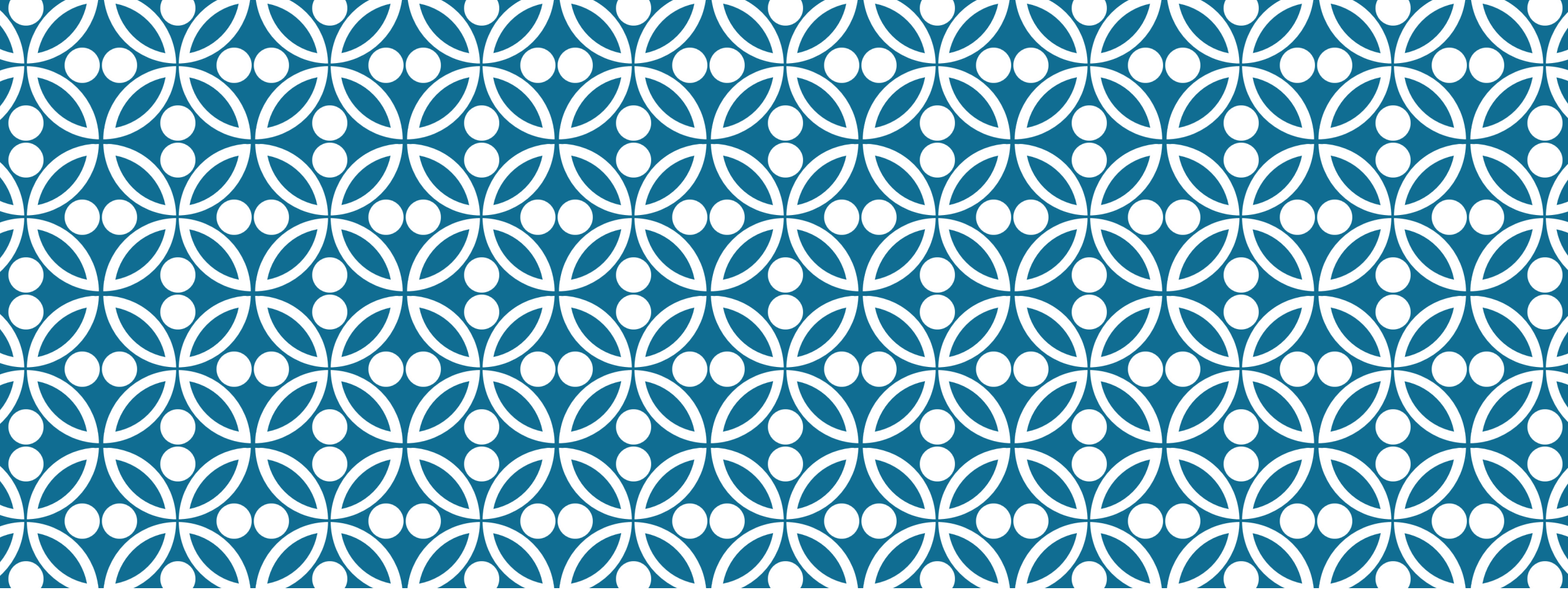


INTELIGÊNCIA ARTIFICIAL NA PRÁTICA

Allan Inocêncio de Souza Costa



INTRODUÇÃO

Onde são apresentados conceitos gerais de Aprendizado de Máquina.

APRENDIZADO DE MÁQUINA

Aprendizado de Máquina (Machine Learning) é uma sub-área da IA na qual são estudados algoritmos que aprendem com os dados.

Utiliza apenas raciocínio indutivo.

É a parte da IA mais utilizada na prática, de modo que falaremos somente sobre ela no restante desta apresentação.

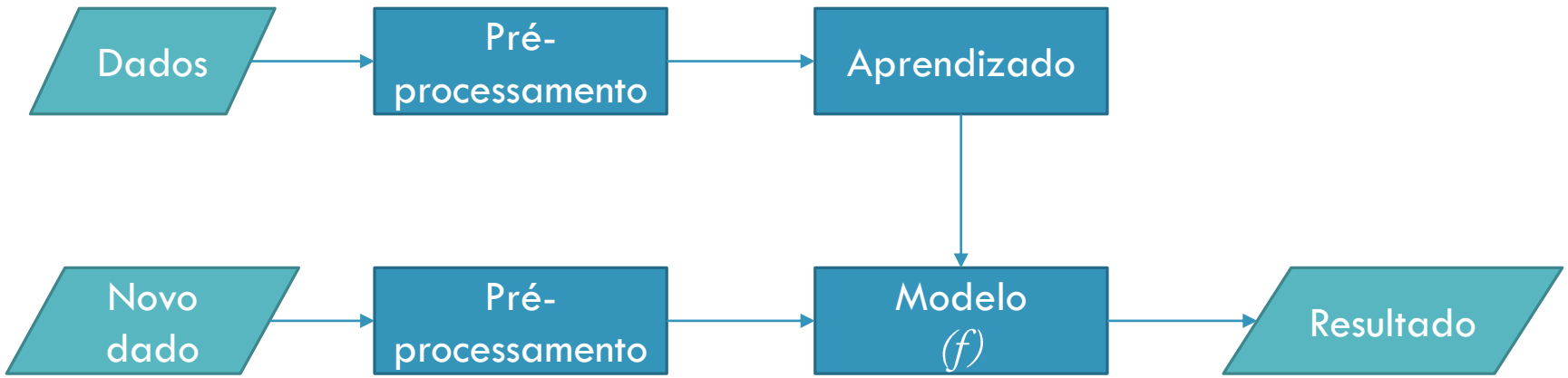
UMA DEFINIÇÃO MAIS FORMAL

"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E"

Mitchell, T. (1997). *Machine Learning*, McGraw Hill. ISBN 0-07-042807-7, p.2.

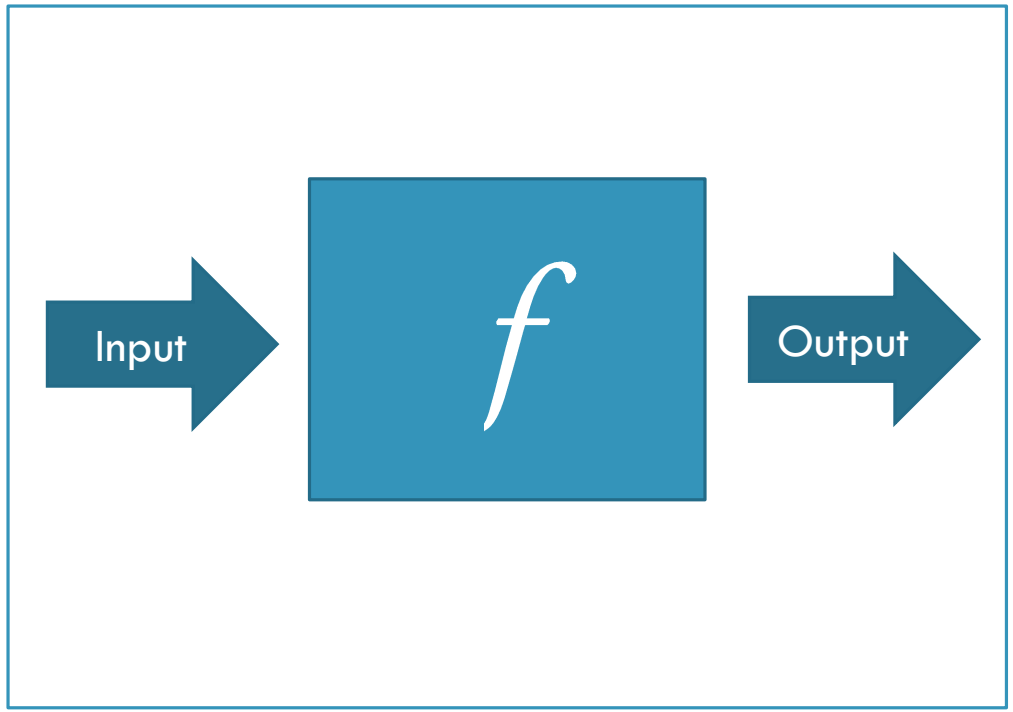
O QUE O ALGORITMO DE FATO APRENDE

FLUXO TÍPICO

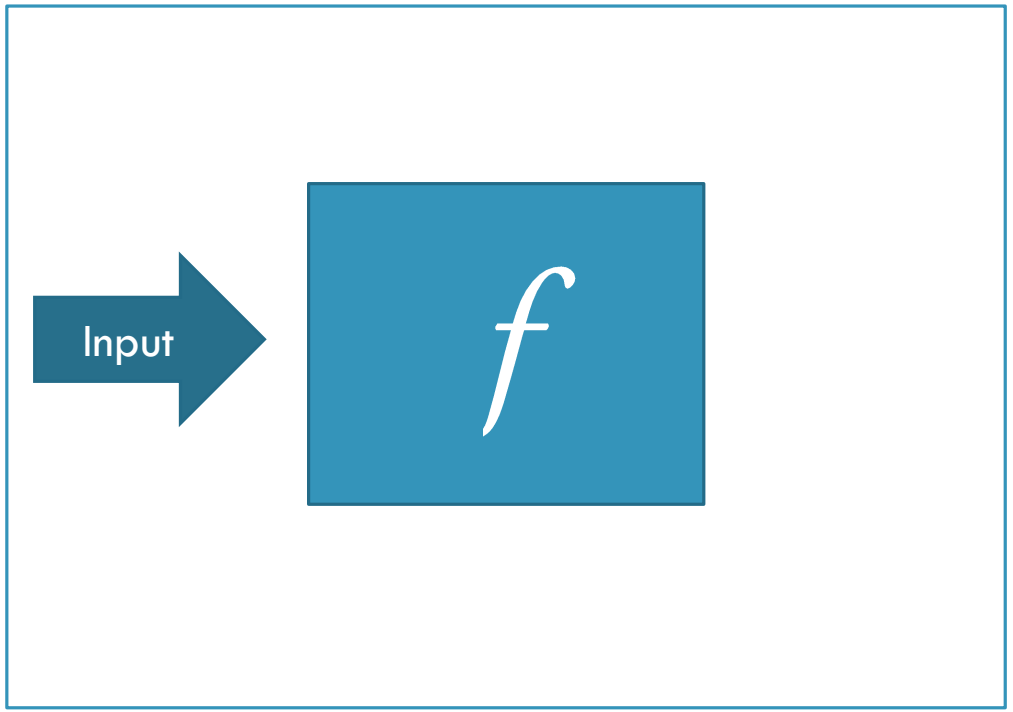


TIPOS DE APRENDIZADO

Aprendizado Supervisionado

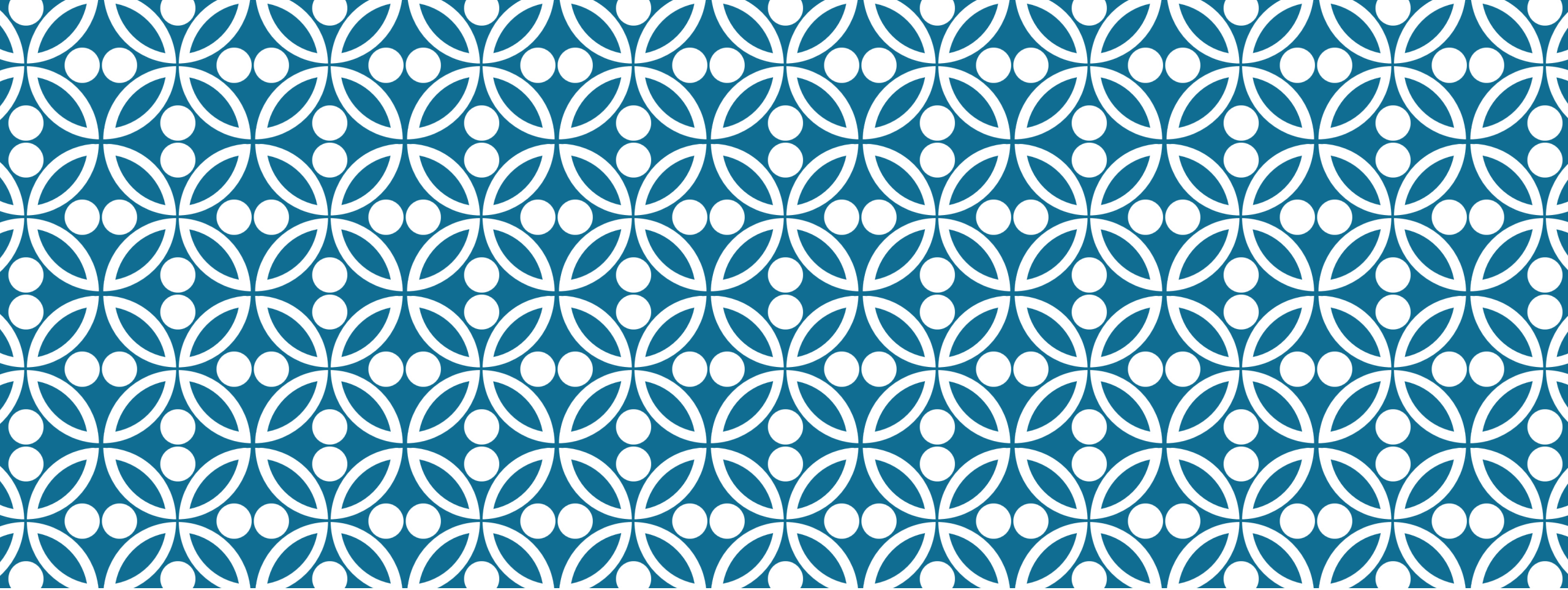


Aprendizado Não-Supervisionado



NOMES DE ALGUNS ALGORITMOS

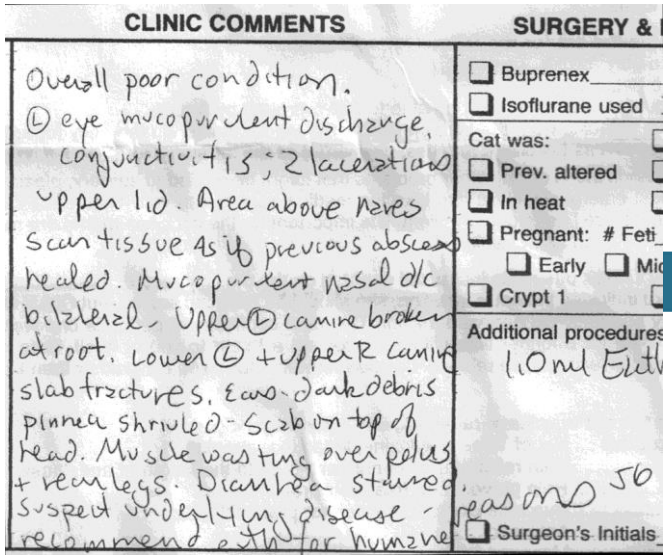
- K-Nearest Neighbors
- Decision Tree
- Random Forest
- Support Vector Machines (SVM)
- Genetic Algorithms
- Neural Networks
- K-Means
- DBSCAN



EXEMPLOS

Onde mostramos uma variedade de problemas que se encaixam na definição de Aprendizado de Máquina.

DIAGNOSTICAR INSUFICIÊNCIA CARDÍACA



https://www.additiveanalytics.com/blog/new-research-predicts-heart-failure-happens/

PREVER O PREÇO DO BITCOIN

Date	24h Average
11/2/2014	826.25
11/1/2014	895.43
10/31/2014	808.89
10/30/2014	917.04
10/29/2014	936.84
10/28/2014	773.1
10/27/2014	944.24
10/26/2014	716.91
10/25/2014	751.57
10/24/2014	757.3

http://newsoffice.mit.edu/2014/mit-computer-scientists-can-predict-price-bitcoin

http://arxiv-web3.library.cornell.edu/pdf/1410.1231v1.pdf

TRADUÇÃO AUTOMÁTICA

https://code.google.com/p/word2vec/

http://arxiv.org/pdf/1309.4168.pdf

CLASSIFICAR TIPOS DE VINHOS

https://archive.ics.uci.edu/ml/datasets/Wine

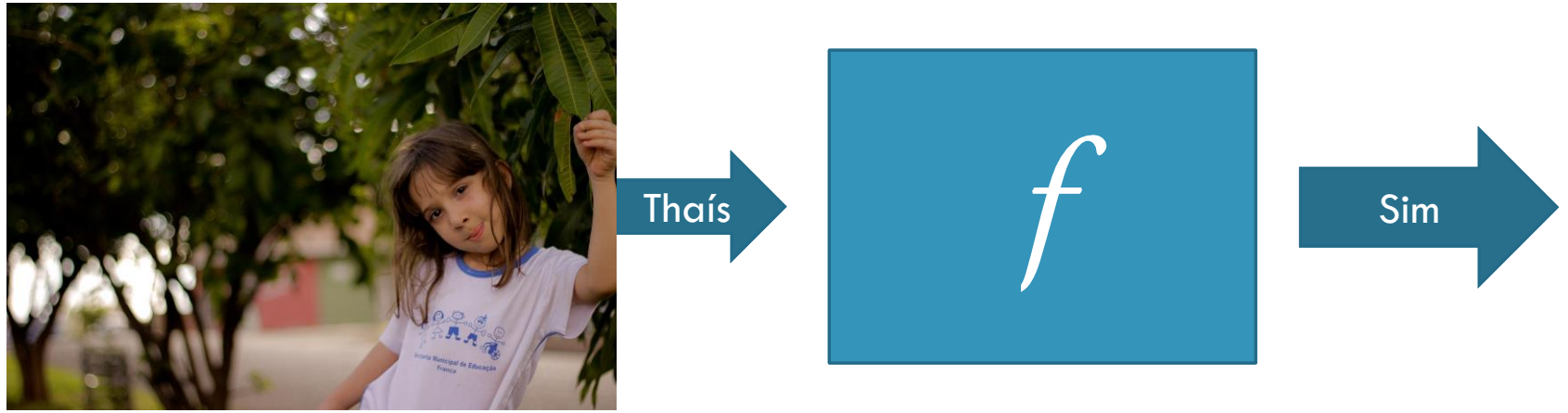
http://mitm613.blogspot.com.br/p/wine-classification.html

DETECÇÃO FACIAL (FACE DETECTION)

https://realpython.com/blog/python/face-recognition-with-python/

http://icml.cc/2012/papers/73.pdf (Google's cat detector)

RECONHECIMENTO FACIAL (FACE RECOGNITON)



http://docs.opencv.org/modules/contrib/doc/facerec/facerec_tutorial.html

https://medium.com/the-physics-arxiv-blog/the-face-recognition-algorithm-that-finally-outperforms-humans-2c567adbf7fc

DETECÇÃO DE OBJETOS (OBJECT DETECTION)

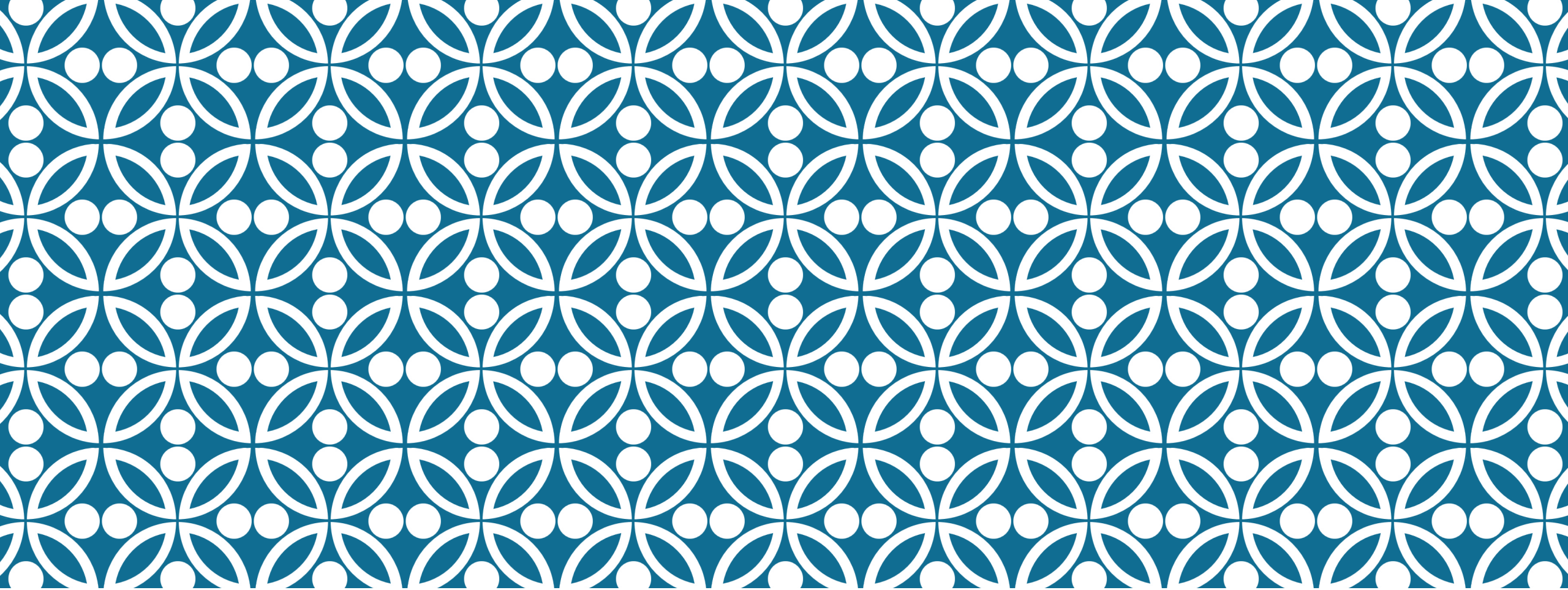
http://tutorial.caffe.berkeleyvision.org/eccv14-diy-pt2.pdf

DETECÇÃO DE ANOMALIAS EM SERVIDORES

https://github.com/cloudwalkio/omg-monitor/

DETECÇÃO DE FRAUDES EM TRANSAÇÕES

https://vimeo.com/109994675



UM ALGORITMO: DBSCAN

Onde explicamos o funcionamento de um algoritmo simples e poderoso.

UTILIZANDO A DENSIDADE DOS PONTOS PARA CRIAR AGLOMERADOS

DBSCAN: Density Based Spatial Clustering of Applications with Noise.

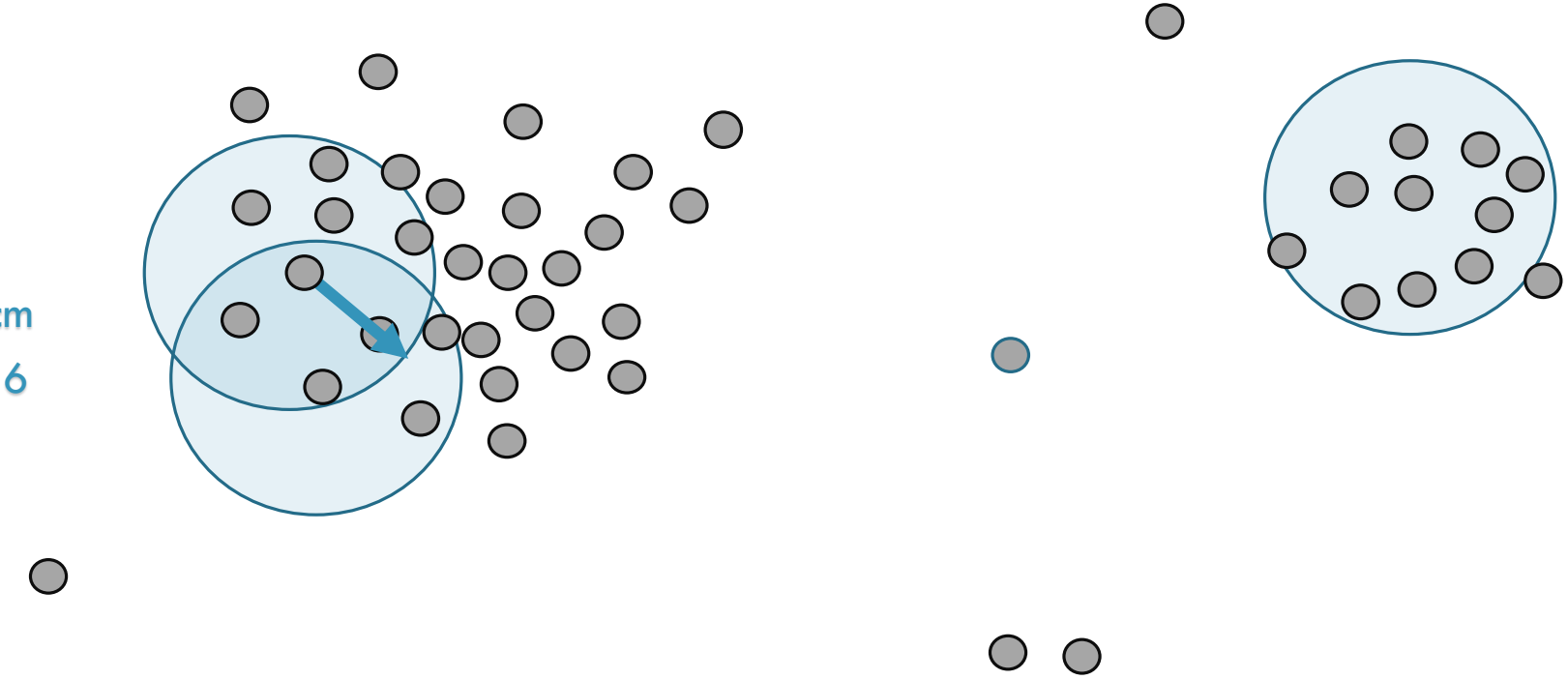
É um algoritmo não-supervisionado cujo objetivo é encontrar aglomerados nos dados.

http://www.dbs.ifi.lmu.de/Publikationen/Papers/KDD-96.final.frame.pdf

A IDEIA BÁSICA

Eps = 2 cm

MinPts = 6

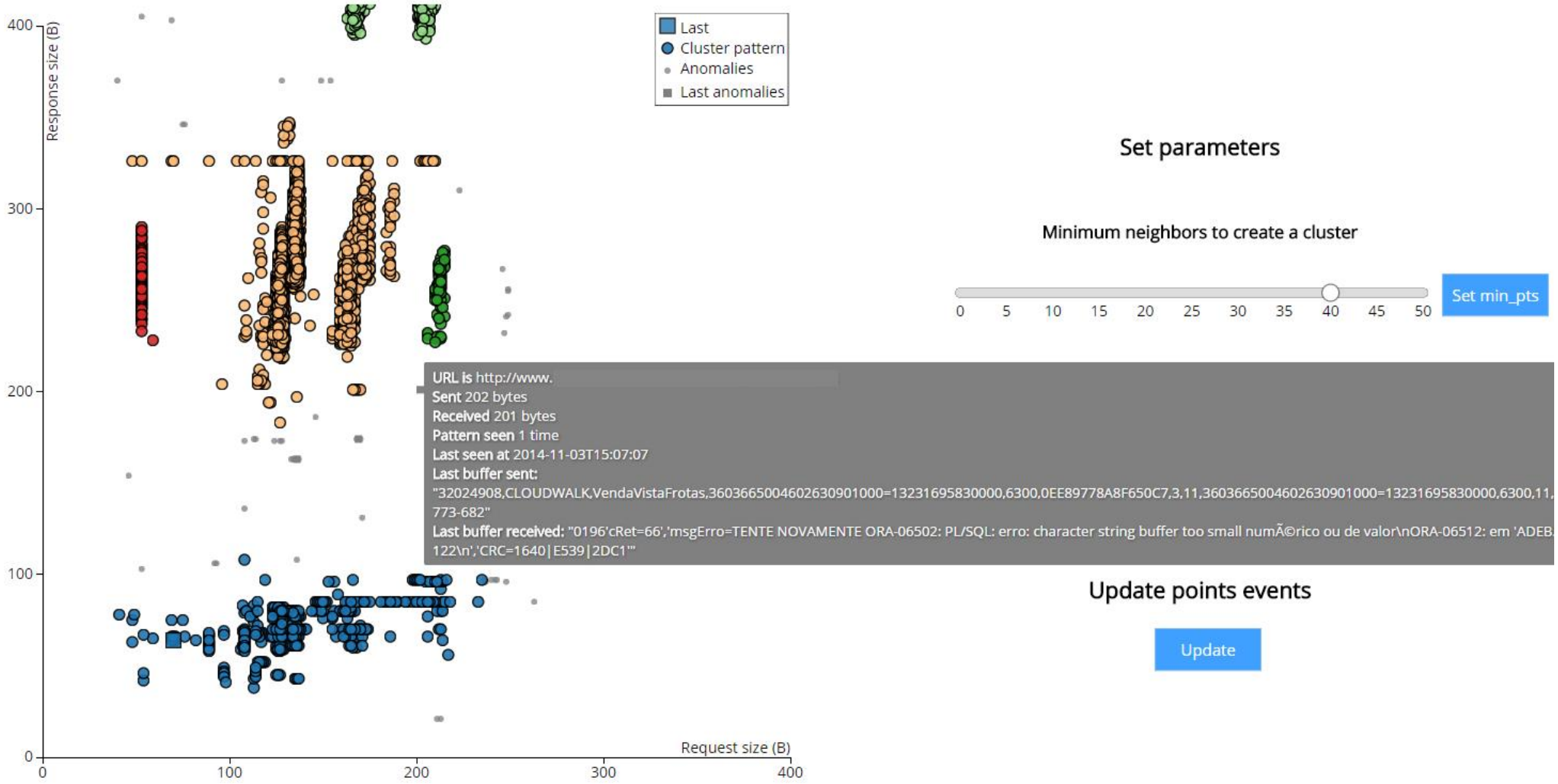


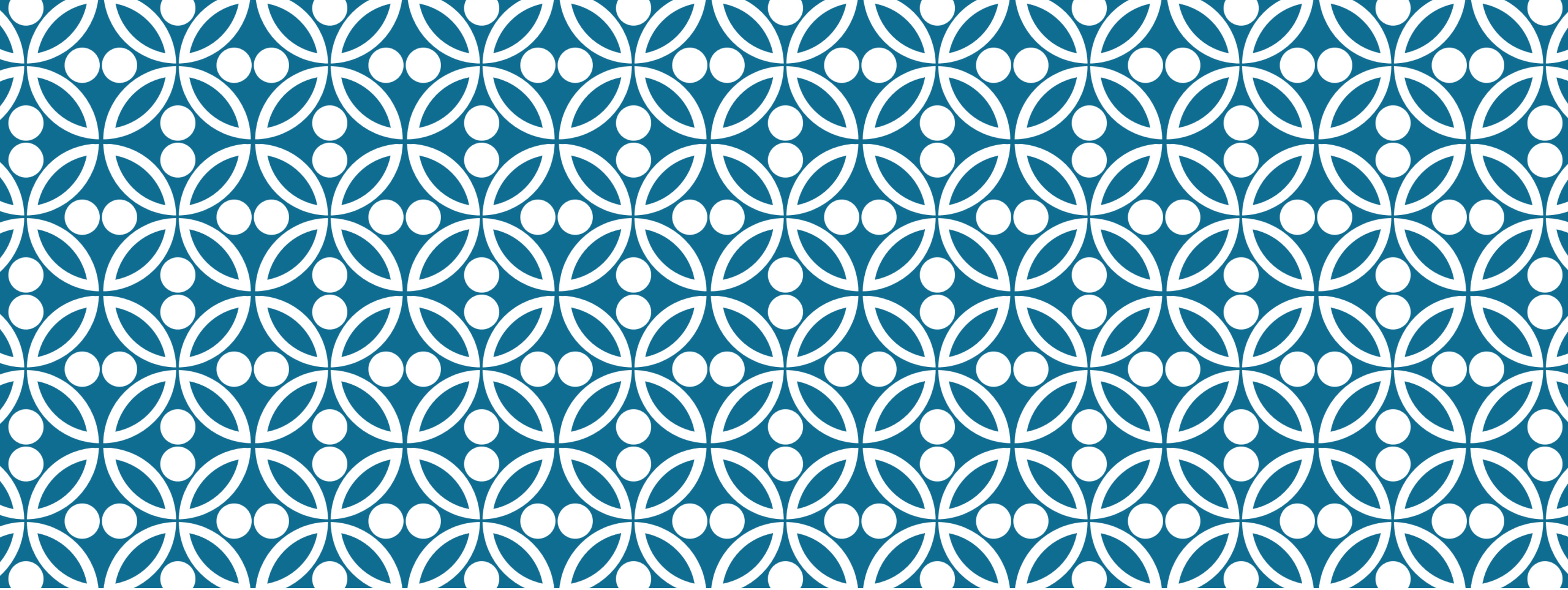
UMA DEMONSTRAÇÃO INTERATIVA

Disponibilizamos um código demonstrando como utilizar nossa implementação do DBSCAN em Python.

Acesse em: http://github.com/allanino/palestra-ia-na-pratica

DETECTANDO ANOMALIAS EM TRANSAÇÕES





DICAS DE ESTUDO

Onde compartilhamos um pouco de experiência com aqueles que querem iniciar os estudos na área.

APRENDA INGLÊS

Há algum material em português, mas o estado-da-arte só é acessível para aqueles capazes de, pelo menos, ler e interpretar textos técnicos em língua inglesa.

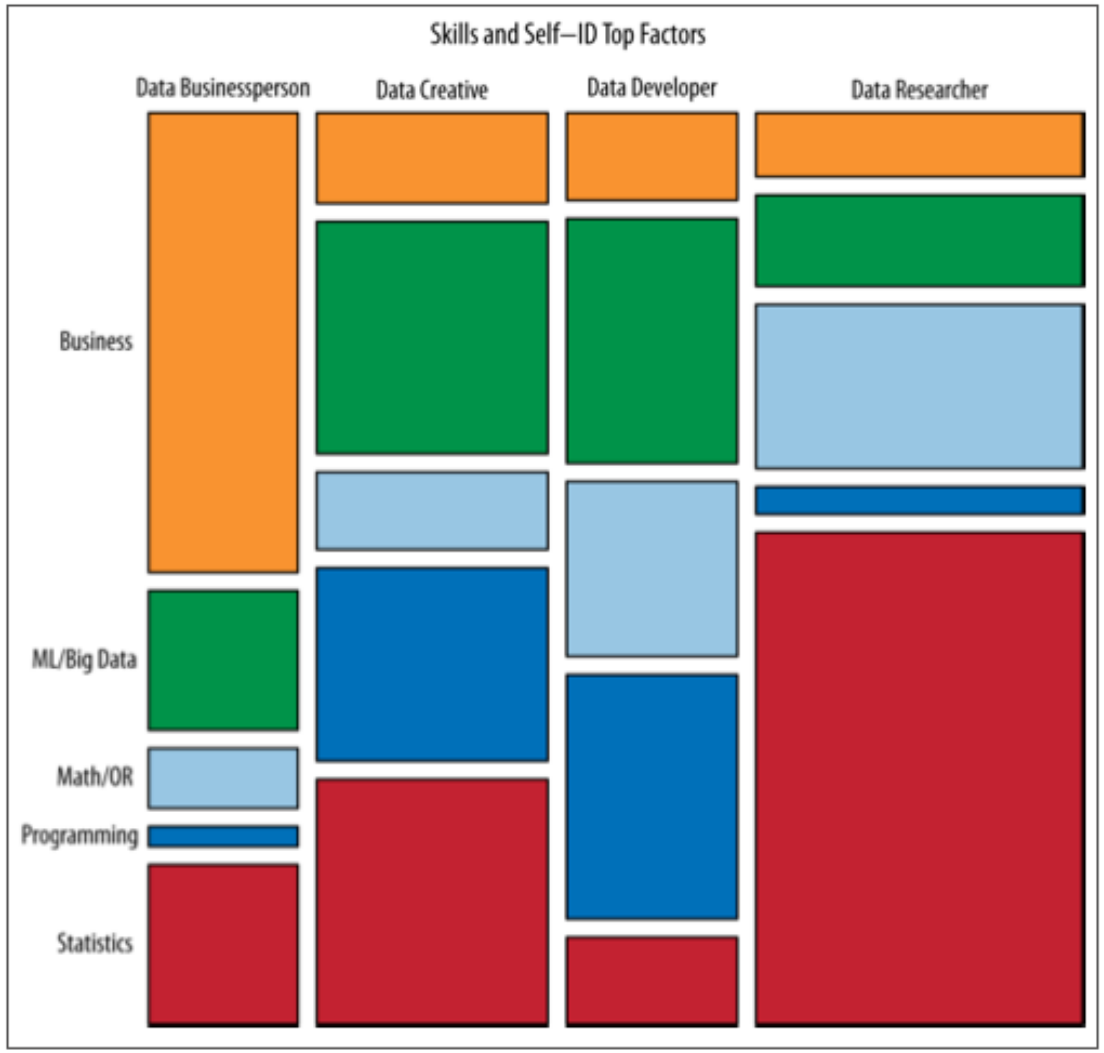
NÃO NEGLIGENCIE A MATEMÁTICA

Não precisa amar Matemática, mas no mínimo respeitá-la e reconhecer seu valor.

Três matérias em particular merecem atenção:

- Cálculo Diferencial
- Estatística
- Programação Linear

QUATRO TIPOS DE CIENTISTAS DE DADOS



Data Businesspeople are the product and profit-focused data scientists. They're leaders, managers, and entrepreneurs, but with a technical bent. A common educational path is an engineering degree paired with an MBA.

Data Creatives are eclectic jacks-of-all-trades, able to work with a broad range of data and tools. They may think of themselves as artists or hackers, and excel at visualization and open source technologies.

Data Developers are focused on writing software to do analytic, statistical, and machine learning tasks, often in production environments. They often have computer science degrees, and often work with so-called "big data".

Data Researchers apply their scientific training, and the tools and techniques they learned in academia, to organizational data. They may have PhDs, and their creative applications of mathematical tools yields valuable insights and products.

http://radar.oreilly.com/2013/06/theres-more-than-one-kind-of-data-scientist.html

INSCREVA-SE NO COURSERA

Excelentes cursos na área.

Em particular, faça o curso de Aprendizado de Máquina de Stanford, ministrado pelo professor Andrew Ng:

https://www.coursera.org/course/ml

Se gostar de Matemática, dê uma olhada no curso "Coding the Matrix: Linear Algebra through Computer Science Applications":

https://www.coursera.org/course/matrix

ESTUDE SOBRE DEEP LEARNING

http://ufldl.stanford.edu/tutorial/

http://neuralnetworksanddeeplearning.com/about.html (em andamento, muito bom material)

OBRIGADO!

allan@allanino.me

http://allanino.me/

http://about.me/allanino/

http://github.com/allanino/

https://www.flickr.com/allanino/

https://www.facebook.com/allan.i.costa/

https://twitter.com/allanino/