



De Analista à Cientista

Jessica Cabral Carvalho
Cientista de Dados
jessica_5548@hotmail.com
11 94819-6501

 /jessica-cabral-carvalho/

 /jcabralc

<https://jcabralc.wordpress.com>

Fevereiro 2019



Quem sou eu

- Jessica Cabral Carvalho – 24 anos
- Técnico de Informática pela ETEC de Ferraz de Vasconcelos
- Graduação pela Fatec Mogi das Cruzes – Julho/16
- Especializações
 - Formação Cientista de Dados
 - Formação Inteligência Artificial
 - Formação Análise Estatística para Cientista de Dados
- Cientista de Dados na OMOTOR

Foco: Análise de dados e soluções cognitivas utilizando IA.

ETEC de Ferraz de Vasconcelos

- 1 anos e meio
- Melhor aluna na ETEC de Ferraz de Vasconcelos
- Programa de intercambio Centro Paula Souza – Chicago – 2012

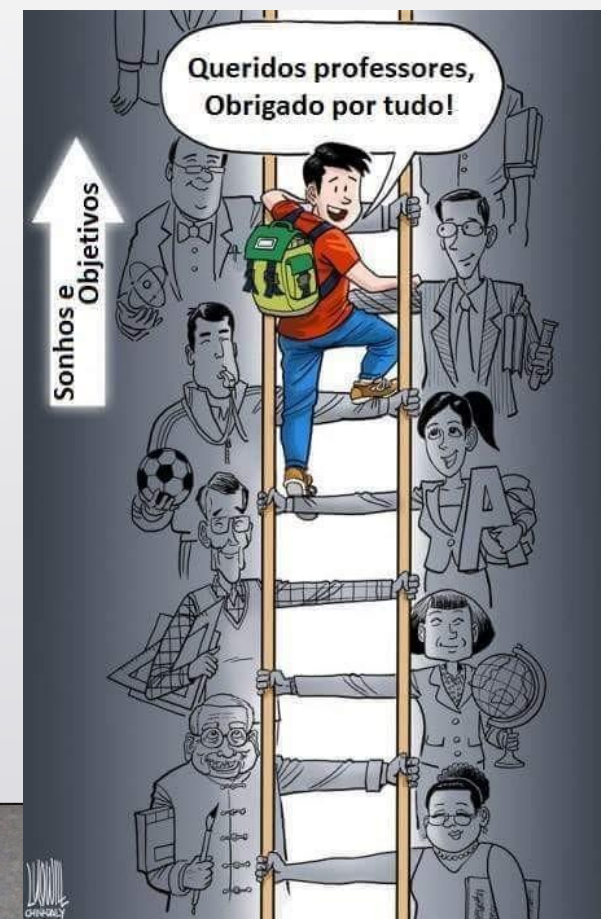


Homem Poder Perfeição



FATEC Mogi das Cruzes

- 4 anos
- Projetos
 - Academia de Muay Thai
 - SCCLIN - TCC
- Estagio na ARSESP
- Analista de Soluções ECM
 - DIPOL – Departamento de Policia de SP – Watson Explorer
 - Solução para identificação de possíveis suspeitos



OMOTOR

- Empresa focada em inovações baseadas em Inteligência Artificial.
- Mones



Responsabilidades

Coleta, higienização e
enriquecimento de
dados

Machine
Learning

Visão
Computacional

Processamento de
Linguagem Natural



Statistic

Visualisation

Pattern
recognit.



Data
Mining

Machine
Learning

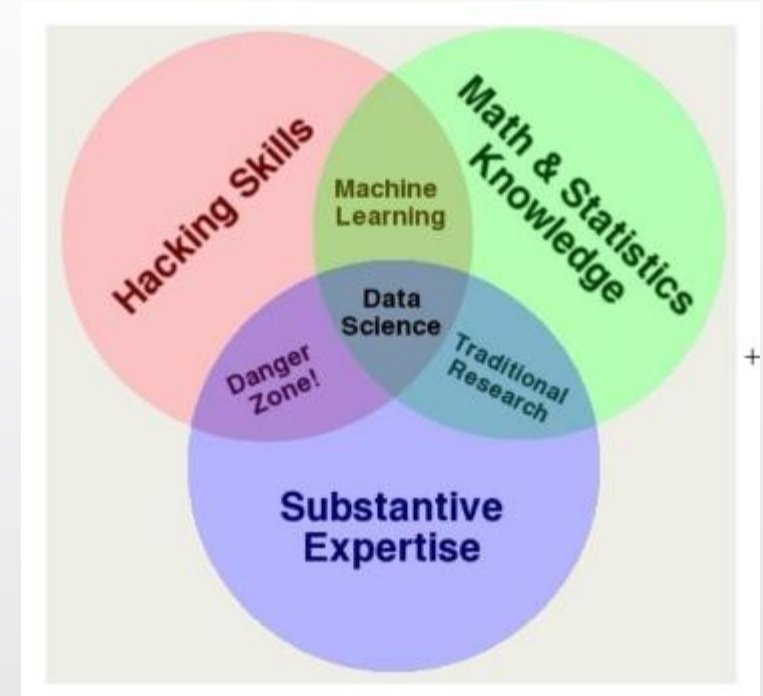
& Database
Data process

Neurocomputing

Ciência de Dados

Cientista de Dados

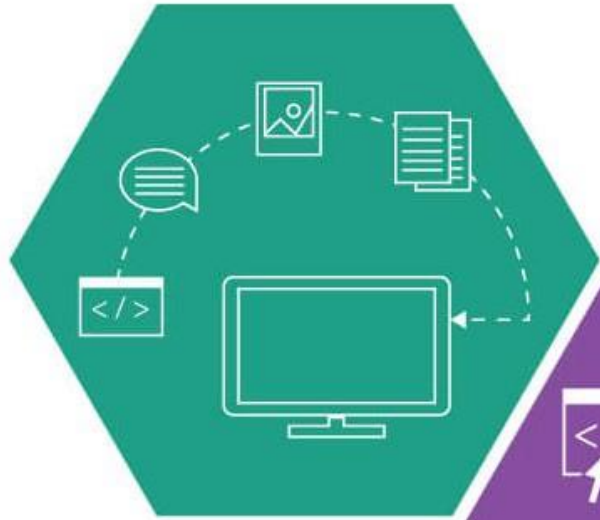
- Área interdisciplinar voltada para o estudo e a análise de **dados** para extração de conhecimento ou insights para possíveis tomadas de decisão, de maneira similar
- Profissional com habilidades técnicas para resolver problemas complexos – e a curiosidade de explorar quais são os problemas que precisam ser resolvidos
- Skills
 - Data mining
 - Machine Learning
 - Python
 - R
 - Data Analysis
 - SQL
 - Algoritmos
 - Entre outros



“Dados são o novo petróleo”

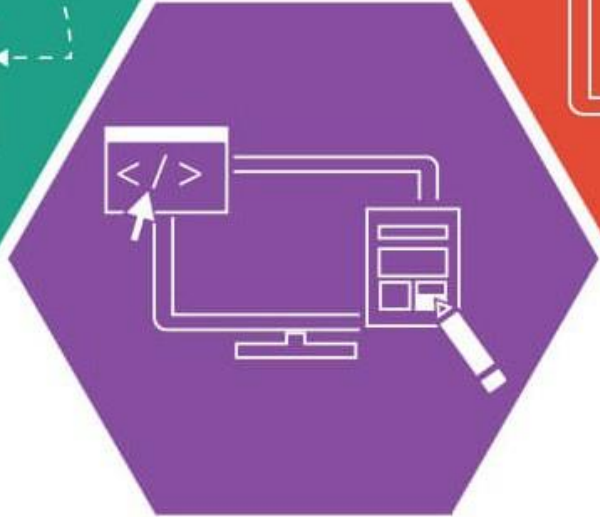


Get Data



1

2



Clean, Prepare
& Manipulate Data

Train Model



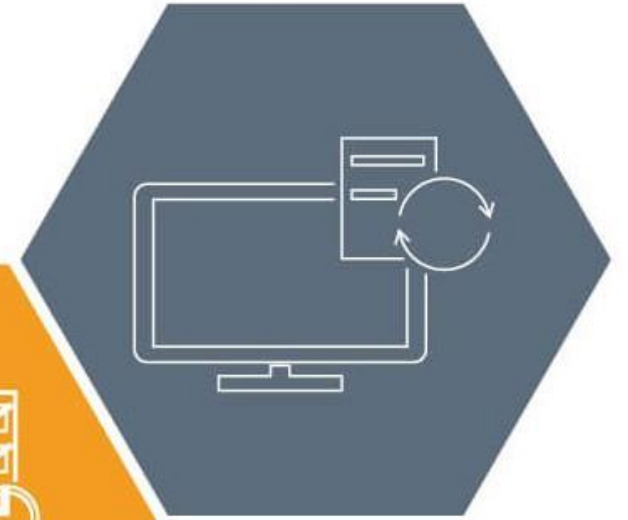
3

4



Test Data

Improve



5

Ferramentas

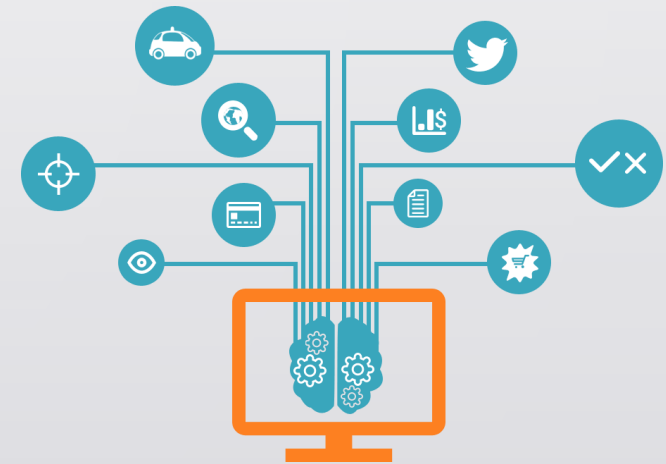




Machine Learning

Machine Learning

- Subcampo da ciência da computação que evoluiu do estudo de reconhecimento de padrões e da teoria do aprendizado computacional em inteligência artificial.
- Basicamente:
 - Um algoritmo que aprende a resolver o problema a partir de vários exemplos.
- Exemplo:
 - Dados varias informações sobre um paciente conseguimos prever se ele irá ou não desenvolver Diabetes





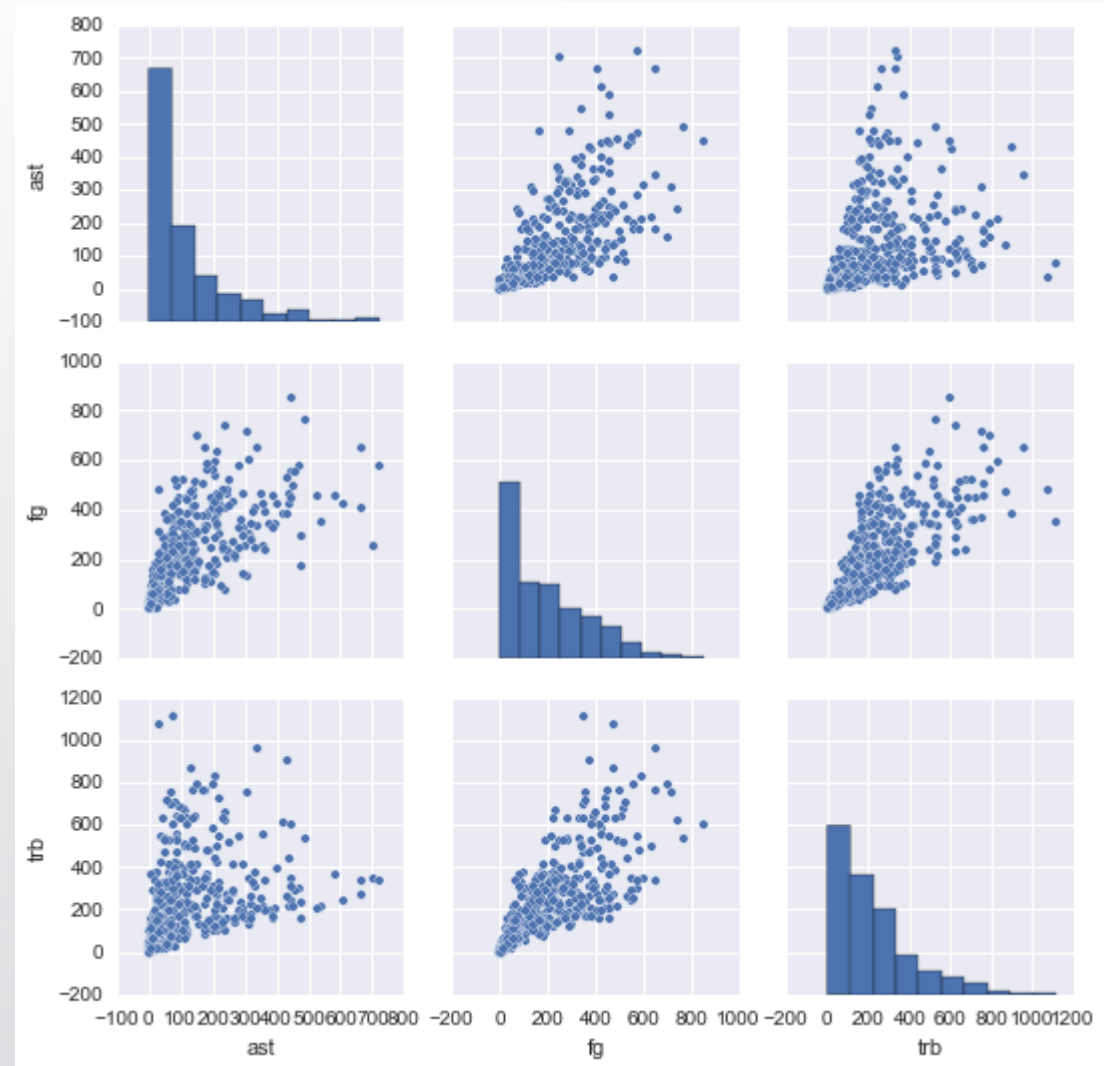
Tipos de Machine Learning

- Supervisionado
 - Classificação
 - Regressão
- Não Supervisionado
 - Clusterização
- Aprendizado por reforço
- Algoritmos de Recomendação

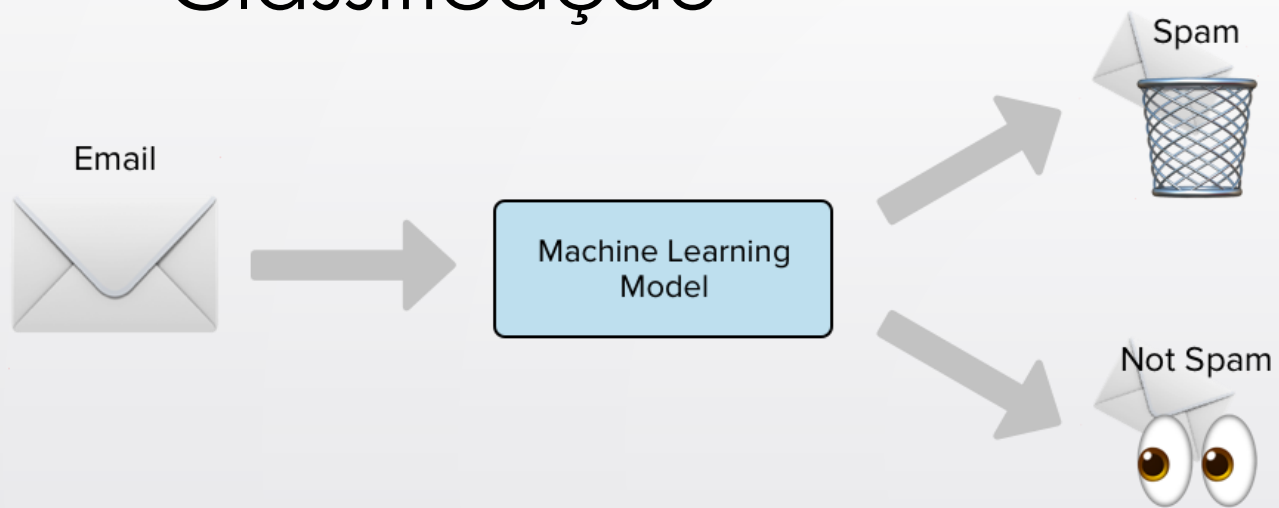
- Visão Computacional
- Processamento de Linguagem Natural
- Modelagem de Tópicos

Áreas de atuação

- Marketing Analytics
- Social Network Analytics
- RH Analytics
- Financial Analytics
- Text Analytics
- Saúde
- Visão Computacional
- ...



Classificação



IRIS dataset



Iris Versicolor



Iris Setosa



Iris Virginica



Relationship

“Dog”

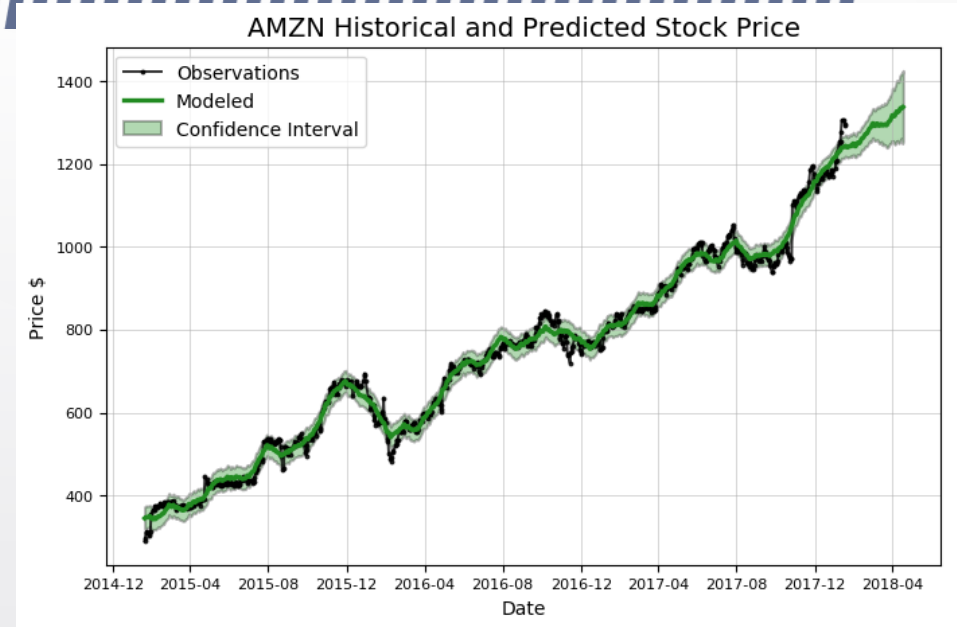


Relationship

“Cat”



Regressão



Clusterização



Identificação Facial



Rostos Encontrados: 21



*** Rosto nº 4 encontrado, confiança: 96.21834564208984

Sexo: Male

Idade entre 17 e 27

Oculos

ANGRY 82.78751373291016

SURPRISED 19.042890548706055

SAD 1.4868736267089844

*** Rosto nº 5 encontrado, confiança: 99.98120880126953

Sexo: Male

Idade entre 48 e 68

Barba

Bigode

Oculos

Oculos de Sol

SURPRISED 24.829620361328125

HAPPY 12.717485427856445

ANGRY 9.191572189331055

*** Rosto nº 6 encontrado, confiança: 99.99586486816406

Sexo: Female

Idade entre 26 e 43

SAD 37.54988479614258

HAPPY 19.35491943359375

DISGUSTED 5.755800247192383

Identificação de Objetos



Roof



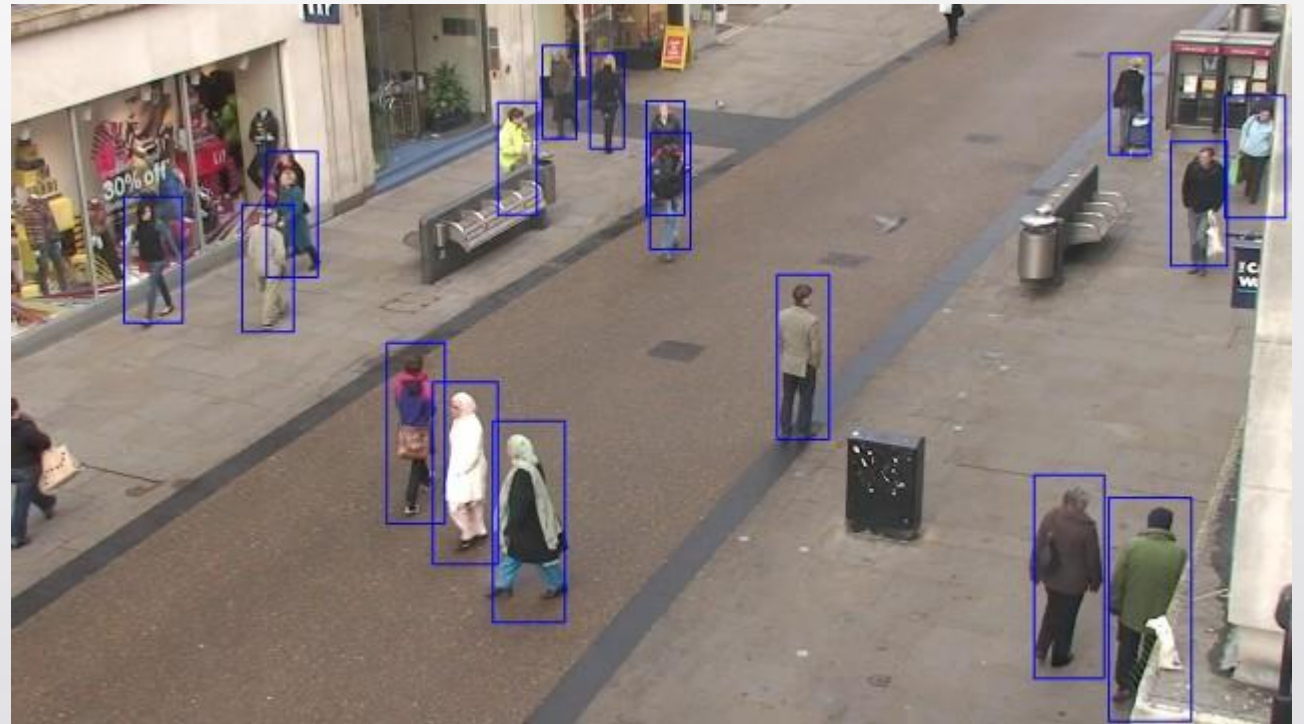
Geyser



Wallplate



Aprone

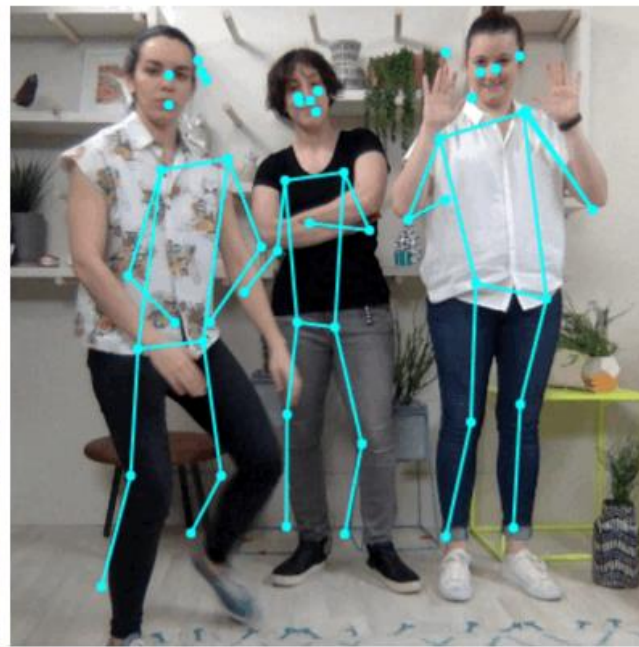
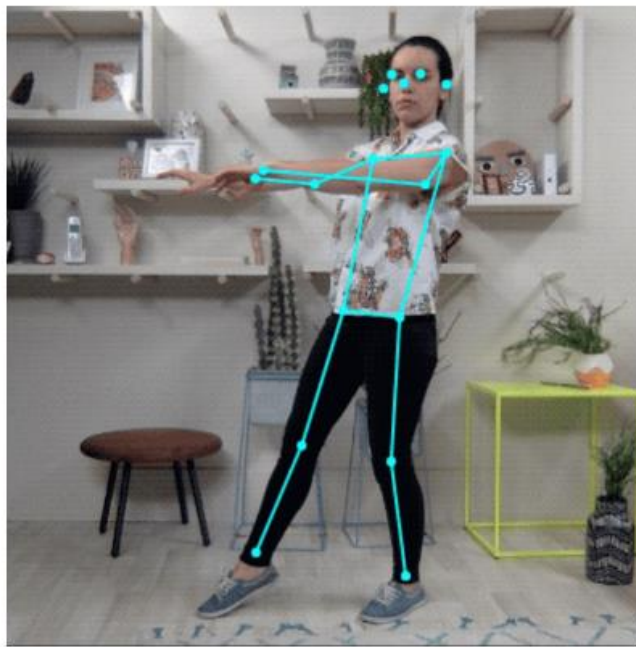


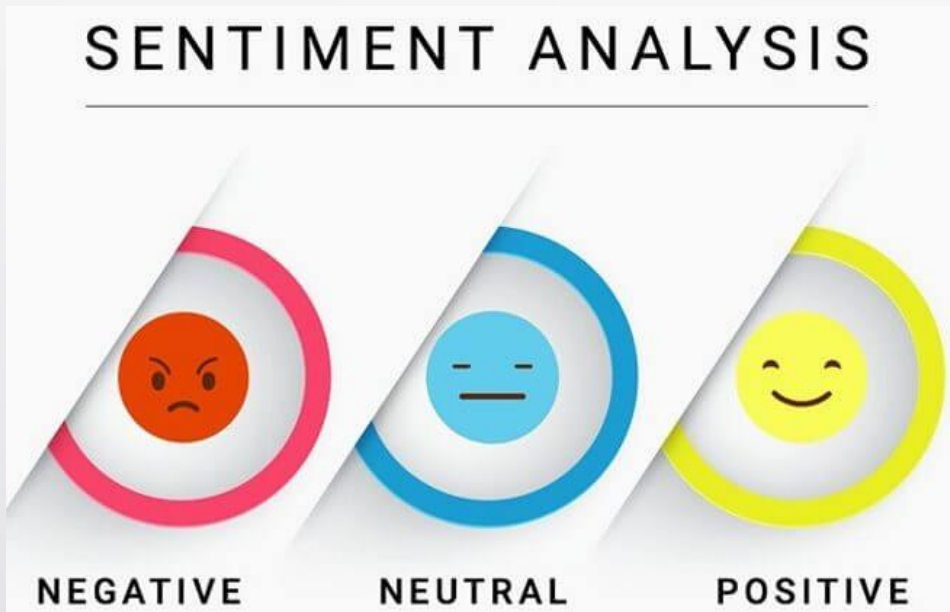
Classificação e caracterização de Objetos



```
Detectando...
{
  "images_processed": 1,
  "images": [
    {
      "classifiers": [
        {
          "classifier_id": "Bebidas_v4_",
          "name": "Bebidas_v4",
          "classes": [
            {
              "score": 0.195159,
              "class": "Jameson"
            },
            {
              "score": 0.0419818,
              "class": "Vinho"
            }
          ]
        }
      ]
    },
    {
      "image": "cropped/image2\\img0.jpg"
    }
  ],
  "custom_classes": 2
}
```


Identificação de Pose



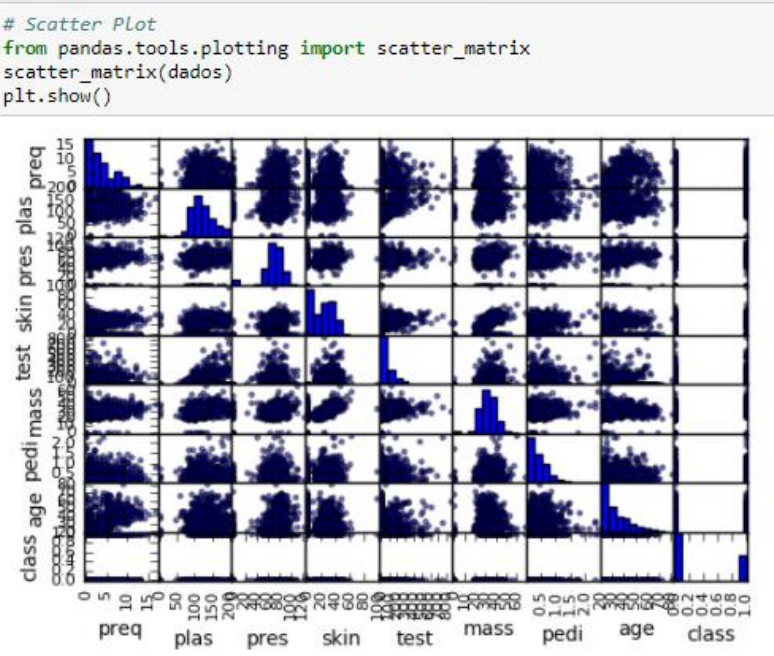


Firing Mr. Strzok PERSON, however, removes a favorite target of Mr. Trump PERSON from the ranks of the F.B.I. GPE and gives Mr. Bowdich PERSON and the F.B.I. GPE director, Christopher A. Wray PERSON, a chance to move beyond the president's ire.

Análise Exploratória de Dados

```
# Correlação de Pearson
dados.corr(method = 'pearson') # Caso n passe nenhum valor, Person é o padrão
```

	preq	plas	pres	skin	test	mass	pedi	age	class
preq	1.000000	0.129459	0.141282	-0.081672	-0.073535	0.017683	-0.033523	0.544341	0.221898
plas	0.129459	1.000000	0.152590	0.057328	0.331357	0.221071	0.137337	0.263514	0.466581
pres	0.141282	0.152590	1.000000	0.207371	0.088933	0.281805	0.041265	0.239528	0.065068
skin	-0.081672	0.057328	0.207371	1.000000	0.436783	0.392573	0.183928	-0.113970	0.074752
test	-0.073535	0.331357	0.088933	0.436783	1.000000	0.197859	0.185071	-0.042163	0.130548
mass	0.017683	0.221071	0.281805	0.392573	0.197859	1.000000	0.140647	0.036242	0.292695
pedi	-0.033523	0.137337	0.041265	0.183928	0.185071	0.140647	1.000000	0.033561	0.173844
age	0.544341	0.263514	0.239528	-0.113970	-0.042163	0.036242	0.033561	1.000000	0.238356
class	0.221898	0.466581	0.065068	0.074752	0.130548	0.292695	0.173844	0.238356	1.000000



Preparação dos Dados

```
# Separando o array em componentes de input e output
X = array[:,0:8]
Y = array[:,8]

# Gerando a nova escala
scaler = MinMaxScaler(feature_range = (0,1))
rescaledX = scaler.fit_transform(X)

# Sumarizando os dados transformados
print(rescaledX[0:5,:])
```

[0.35294118	0.74371859	0.59016393	0.35353535	0.			0.50074516
	0.23441503	0.48333333						
[0.05882353	0.42713568	0.54098361	0.29292929	0.			0.39642325
	0.11656704	0.16666667						
[0.47058824	0.91959799	0.52459016	0.				0.34724292
	0.25362938	0.18333333						
[0.05882353	0.44723618	0.54098361	0.23232323	0.11111111			0.41877794
	0.03800171	0.						
[0.	0.68844221	0.32786885	0.35353535	0.19858156			0.64232489
	0.94363792	0.2						
]								

In [21]: cidadesUF_4

Out[21]:

		CIDADE	POPULAÇÃO	VALOR_TOTAL_PREVISTO	PONTOS_ATENDIDOS	VALOR_INVESTIDO
STATUS	UF					
Sem Previsão	AC	Tarauacá	38819	484.547,53	8	0
	AL	Delmiro Gouveia	51997	1.169.381,70	43	0
	AL	Estrela de Alagoas	18306	527.206,83	39	0
	AM	Autazes	37033	1.171.151,92	51	0
	AM	Benjamin Constant	39484	844.586,48	12	0
	AM	Careiro	36435	748.797,08	17	0
Concluída	AM	Coari	83078	814.954,75	27	0
Sem Previsão	AM	Codajás	26777	757.706,74	36	0
	AM	Iranduba	45984	800.655,02	30	0



Casos de Uso

- 1. Entendendo e direcionando clientes (Marketing)
- 2. Quantificação pessoal e otimização de desempenho (RH)
- 3. Melhorando a saúde e a saúde pública
 - Detecção e prevenção de doenças
- 4. Melhorando o Desempenho Esportivo
 - Filme Moneyball (O homem que mudou o jogo)
- 5. Melhorar a segurança e a aplicação da lei.
- 6. Melhorar e otimizar cidades e países
 - Smart Cities
- 7. Melhoramento e melhor reaproveitamento de recursos ligados a Agro
 - <http://agrodatascience.com/>
- 8. Transporte
 - Previsão de Demanda

Outros exemplos:

<https://www.kdnuggets.com/2017/08/ibm-top-10-machine-learning-use-cases-part1.html>

Aprender data Science pode ser
 muito difícil....



Persistencia é a chave

É sobre **juntar as peças**

Não precisa de um PHD para fazer
data Science

A melhor maneira de **aprender** é
fazendo

Onde Aprender



Data Science
Academy

coursera



UDACITY

Formação
Cientista de
Dados

Formação
Inteligência
Artificial



Johns Hopkins University
Data Science



deeplearning.ai
Deep Learning

IDIOMA:

Fundamentos de Data
Science I

IDIOMA:

Fundamentos de Data
Science II

IDIOMA:

Data Science para
Negócios



Dicas

- Enriqueça seu perfil do LinkedIn (não minta!)
- Github
- Se venda – crie seu portfólio
 - Blog
 - Posts no LinkedIn/Facebook/ Twitter
- Saiba tanto tecnicamente quanto de business
- Nunca, jamais, pare de aprender

Divirta-se!

Obrigada!

