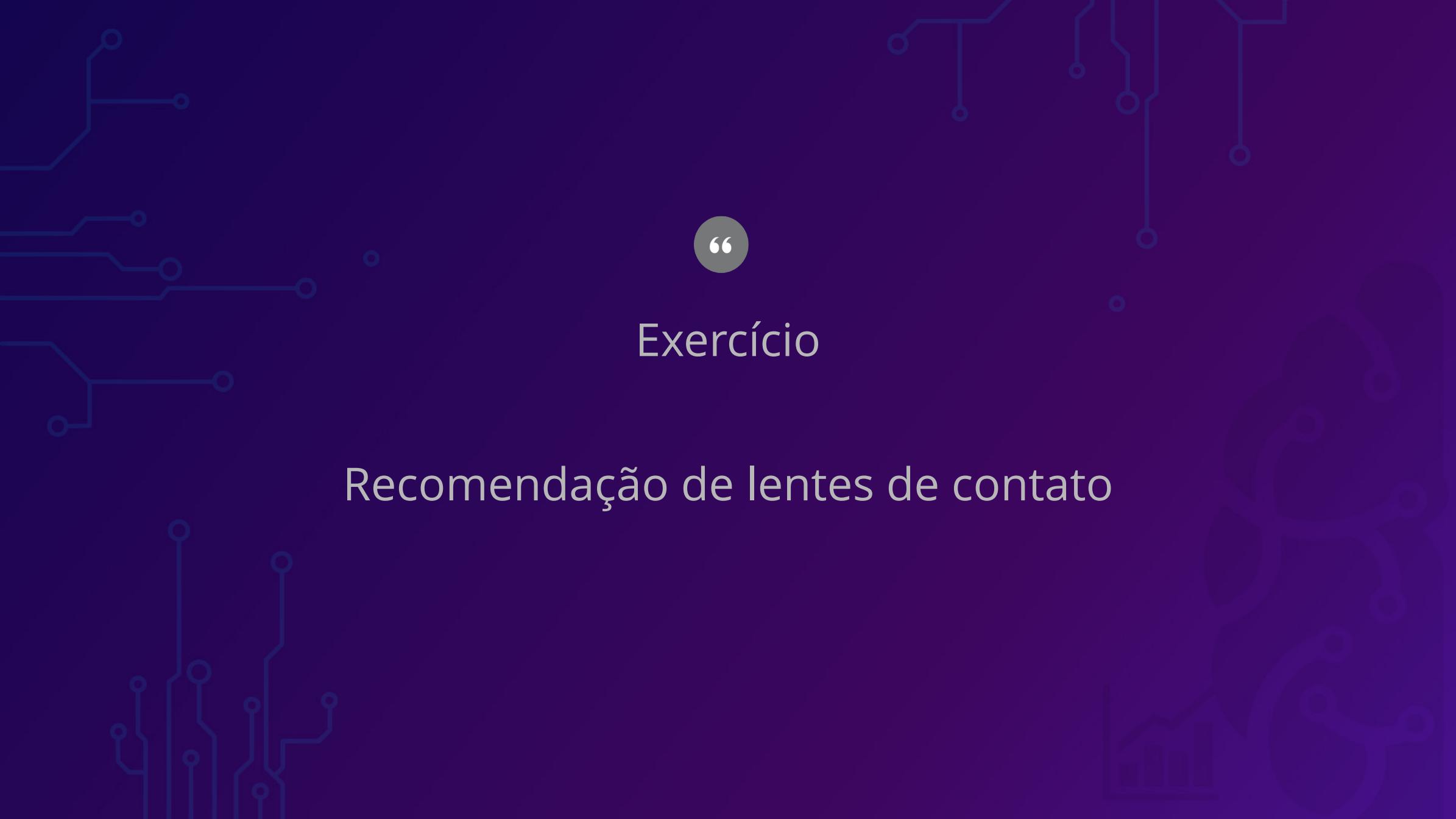


FIAP





Exercício

Recomendação de lentes de contato

- 
- 1** Motivação e **Contextualização Geral**
 - 2** Ciclo de vida do **dado**
 - 3** Ciclo de vida do **modelo**
 - 4** Técnicas e **Algoritmos**
 - 5** Perfil e responsabilidades dos **profissionais de dados**
 - 6** Aplicação e **Cases**

- 
- 1** Motivação e **Contextualização Geral**
 - 2** Ciclo de vida do **dado**
 - 3** Ciclo de vida do **modelo**
 - 4** Técnicas e **Algoritmos**
 - 5** Perfil e responsabilidades dos **profissionais de dados**
 - 6** Aplicação e **Cases**
- 

Data analytics remains **untapped**

69%

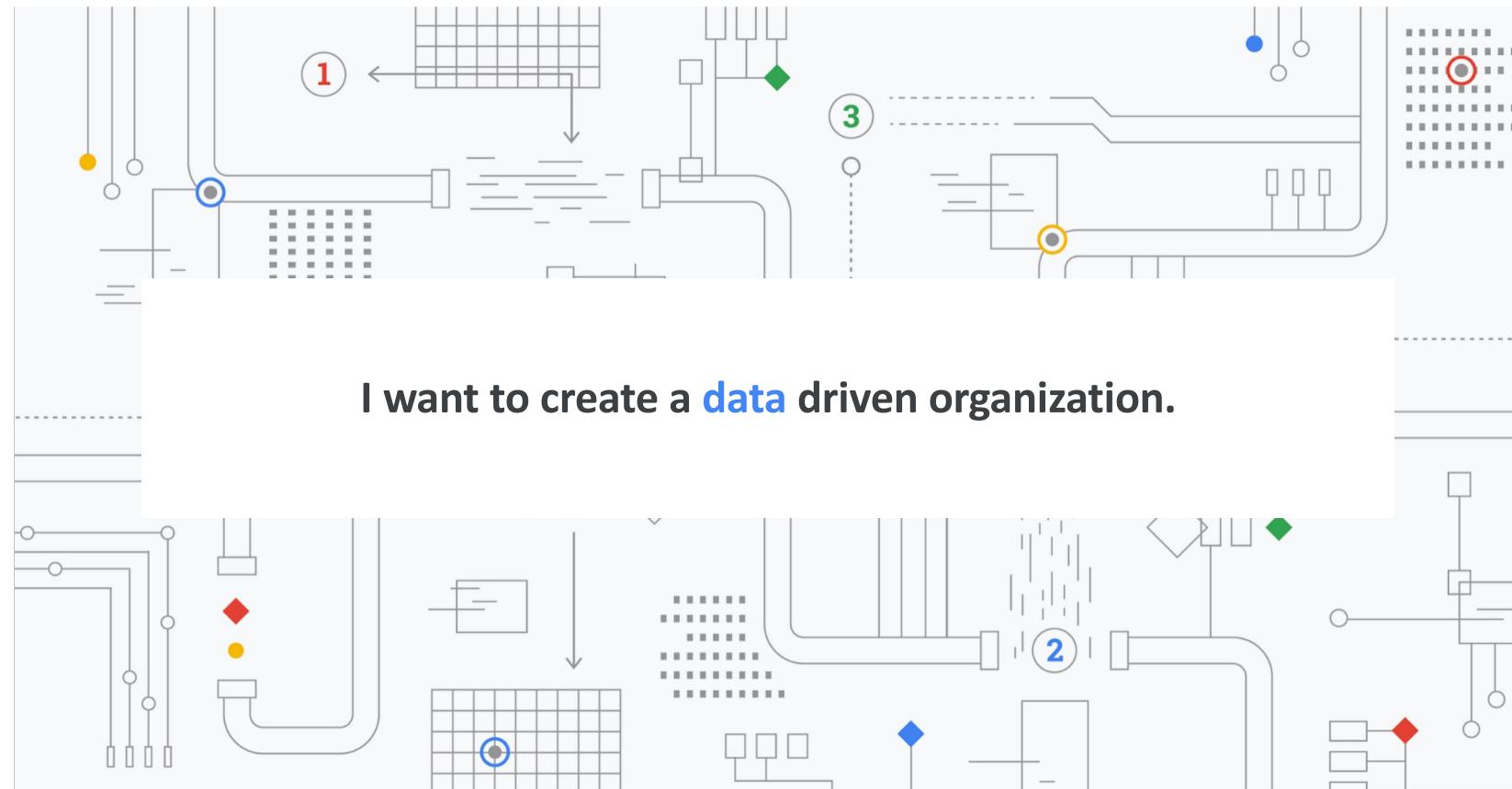
69% of companies report that they have not created a data-driven organization

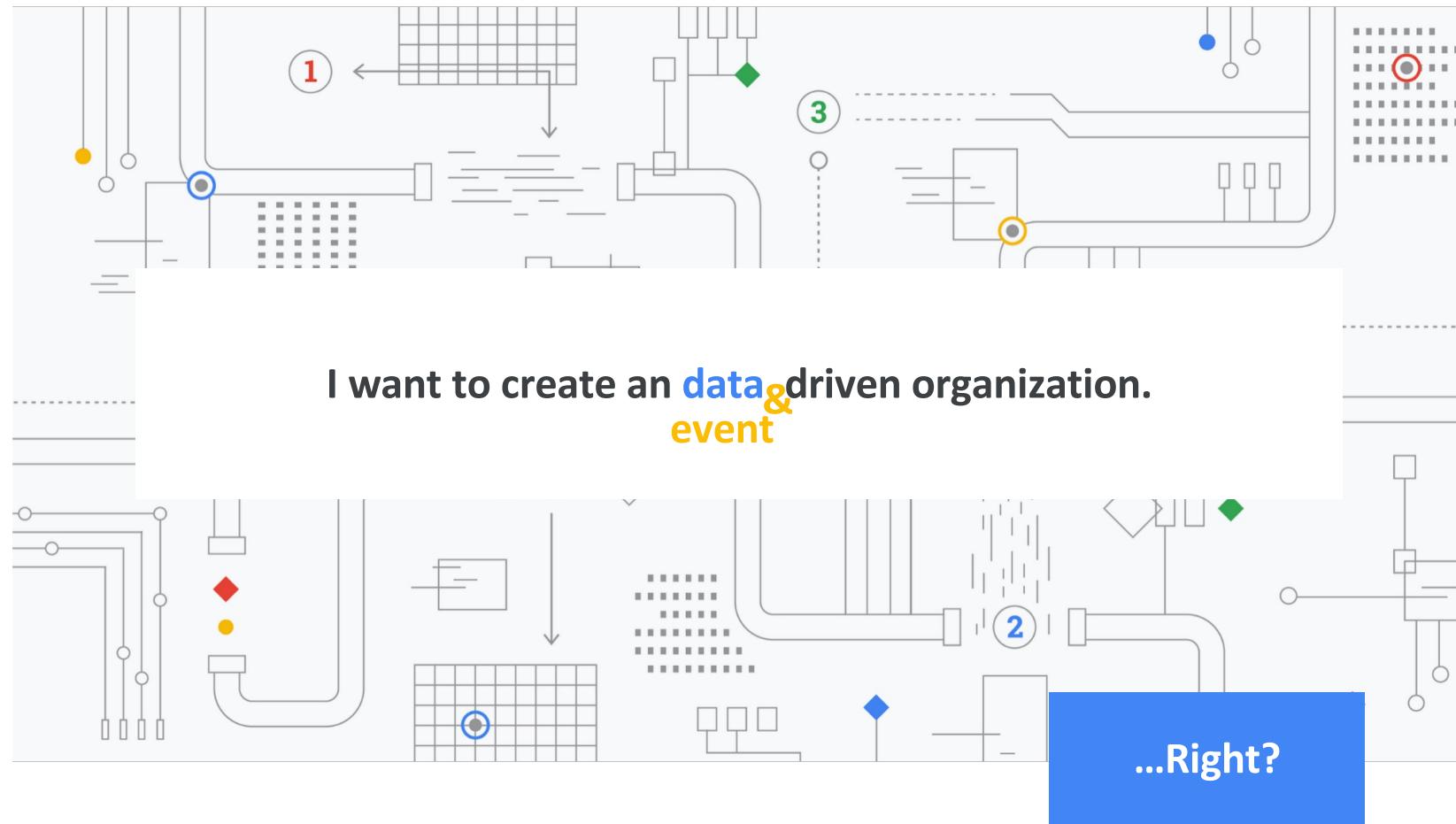
71%

And 71% report that they have yet to forge a data culture

Como as empresas pensam que se tornarão orientadas a dados









**Data-driven organisations use data
for **strategic** purposes**

↑
**Operational success
influences strategic
priorities**

↓
**Strategic vision
defines operational
priorities**



**Event-driven organisations use data
for **operational** purposes**

Common characteristics

Strategic Orientation

- Human-involved
- Develop long-term insights
- Drive quarterly planning
- Determine product strategy
- Determine customer segmentation
- Inform marketing campaigns

Operational Orientation

- Automated processes
- Real-time responsive
- Drive instant interactions
- Determine the product
- Segmentation yields to individual
- Select marketing campaigns

Machine Learning

Machine Learning é uma maneira de obter informações a partir de dados



data



algorithm



insight

“

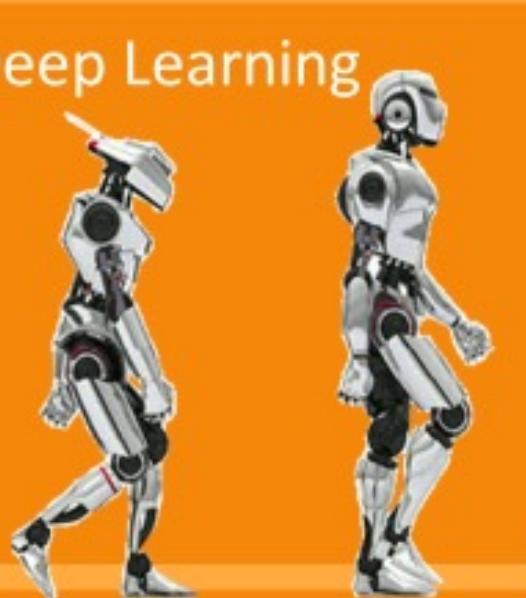
Pode-se dizer que um programa de computador aprende a partir de uma experiência E com respeito a uma classe de tarefas T e com medida de desempenho P, se seu desempenho P na tarefa T melhora com a experiência E.

Mitchell, 1997

Inteligência
Artificial

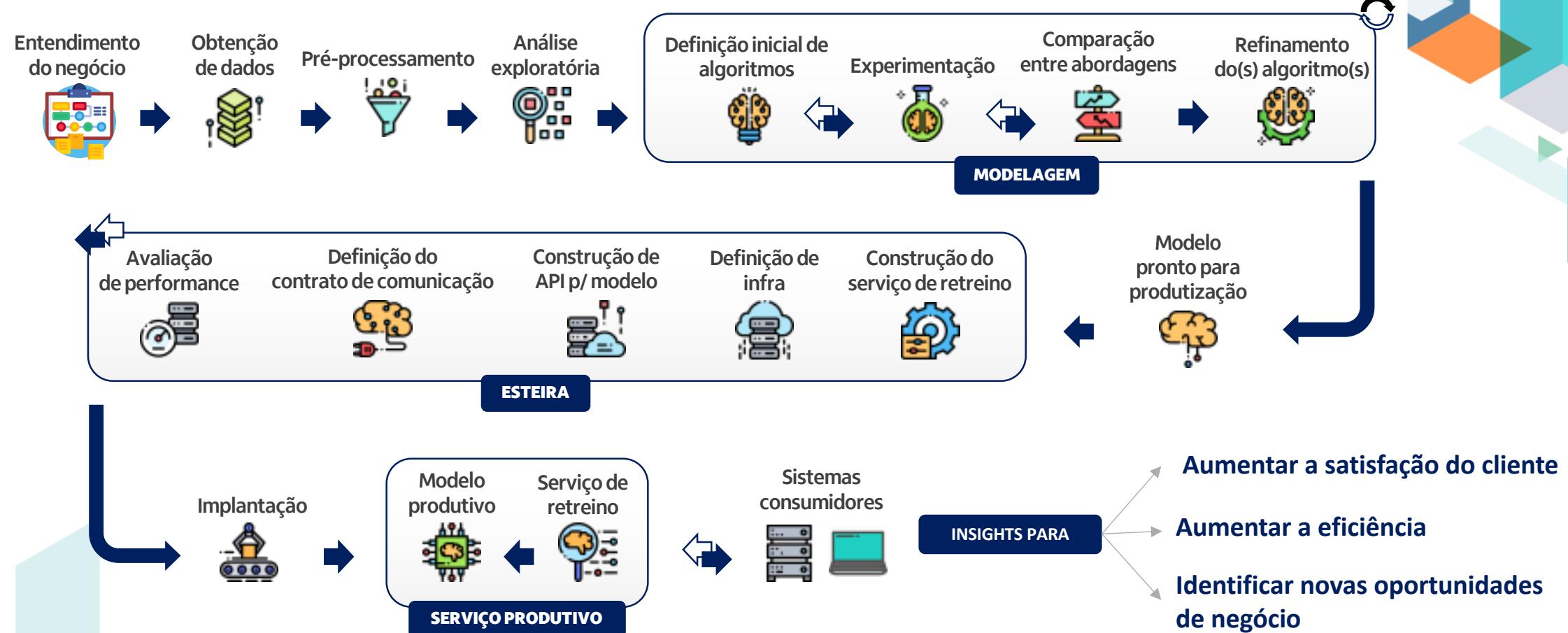
Machine
Learning

Deep Learning

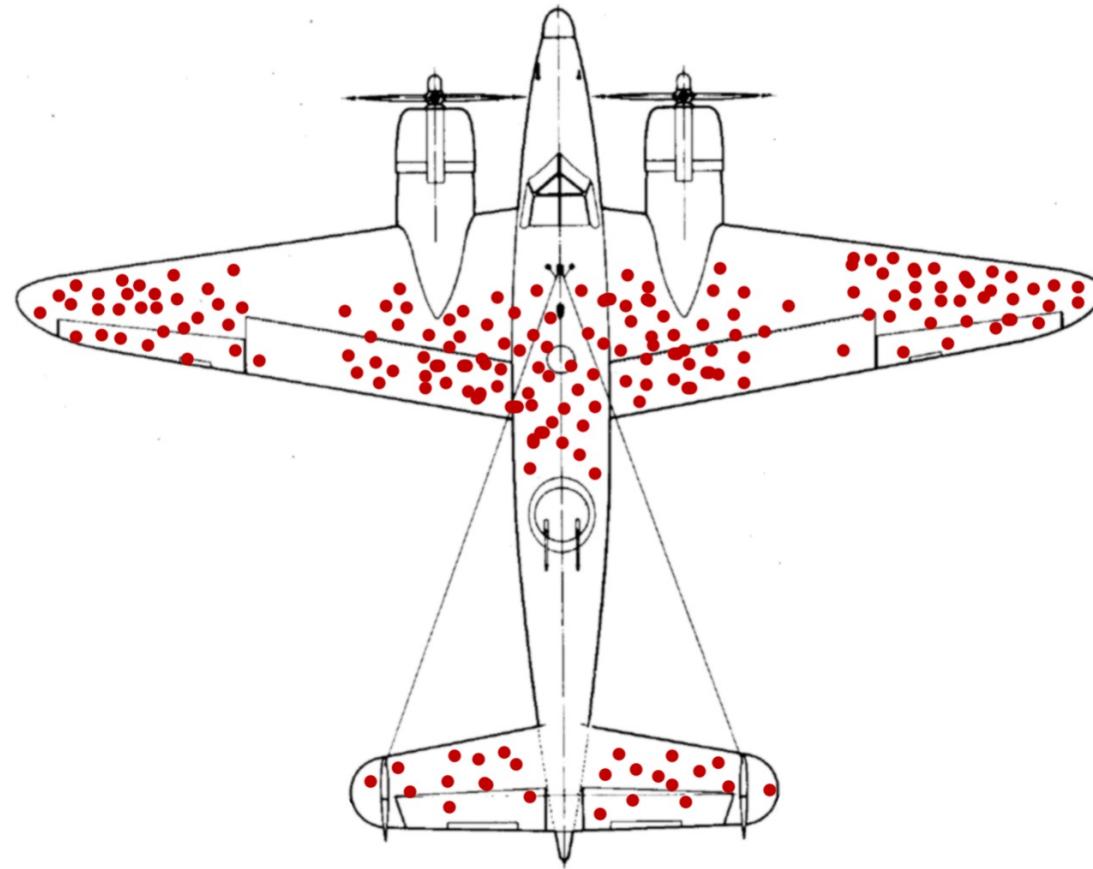


1950 1960 1970 1980 1990 2000 2010 2020

Ciclo de vida de projetos de Dados e Analytics



Entendimento do problema

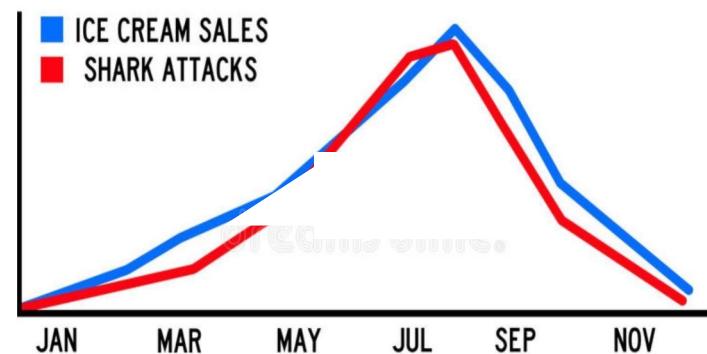


Correlação e consequência

“Em um estudo, constatou-se que o período em que há **mais ataques de tubarão** é entre os meses de março a novembro. Esse também é o período em que **mais se vende sorvete** no país.”

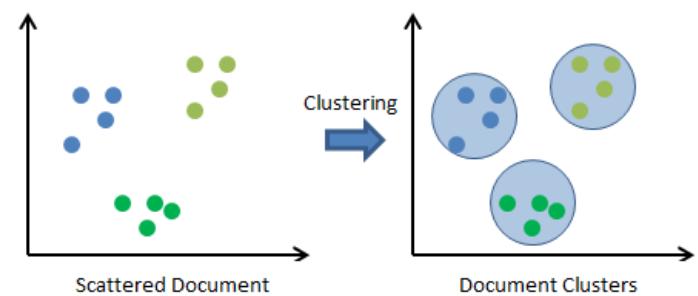
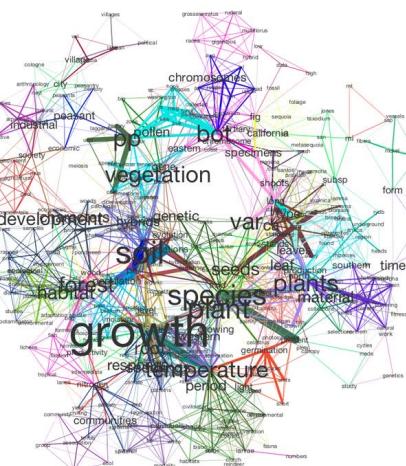
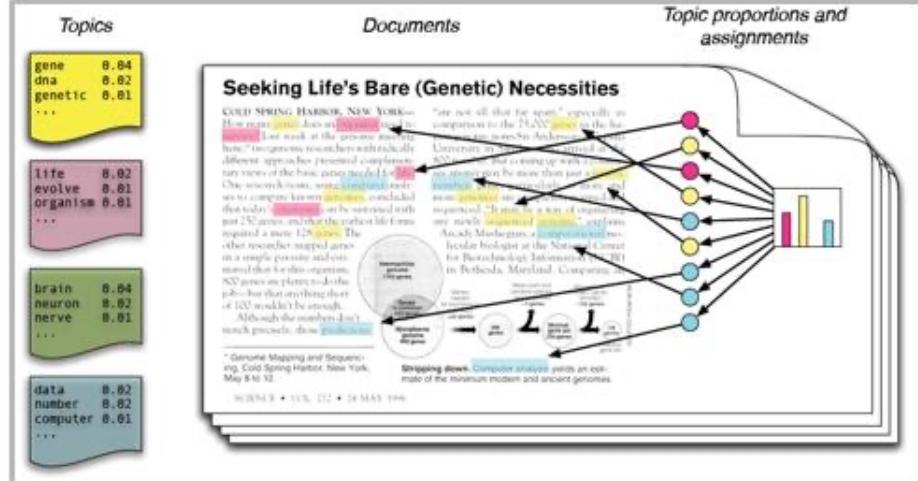
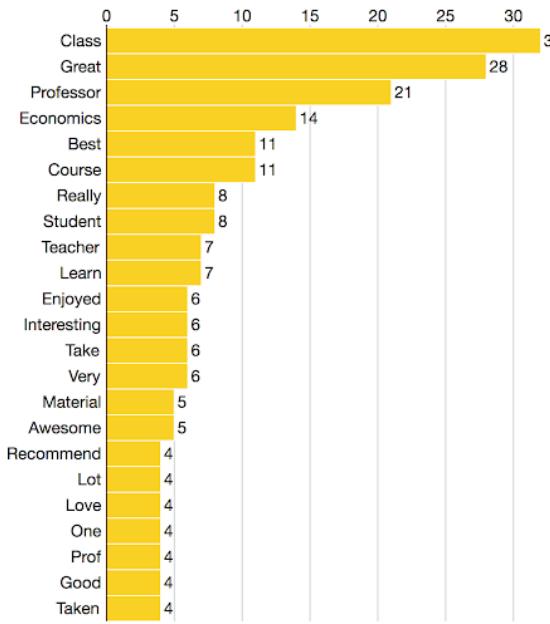
A **venda de sorvete** causa **ataques de tubarão**?

<https://www.tylervigen.com/spurious-correlations>

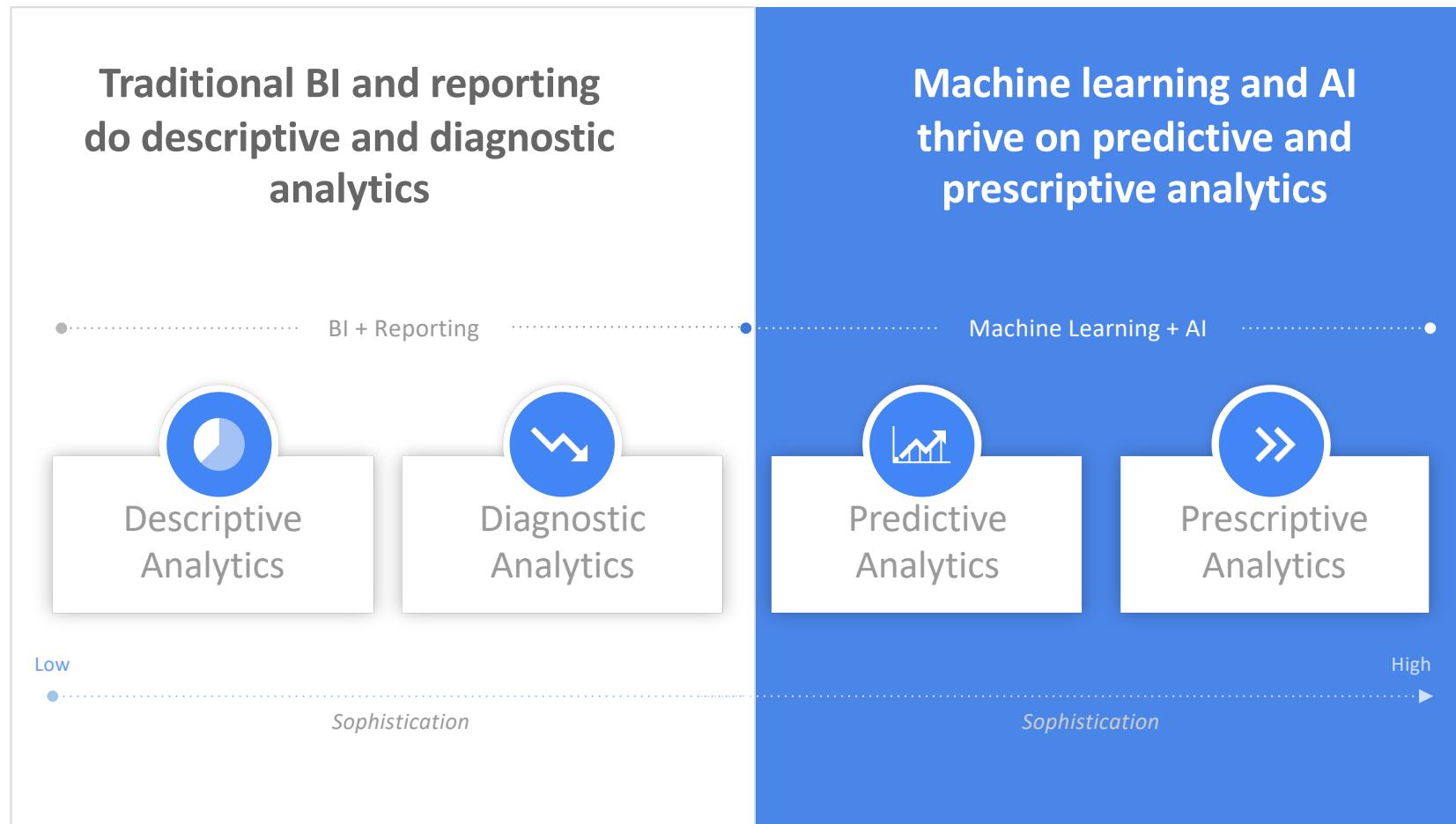


Analytics vs Advanced Analytics

Conversion Marketing



Analytics vs Advanced Analytics



Afinal,



Como funciona inteligência artificial?

Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim
	Luiz	R\$1.500,00	Não
	Maria	R\$5.000,00	Sim

Base de dados inicial

Inteligência Artificial

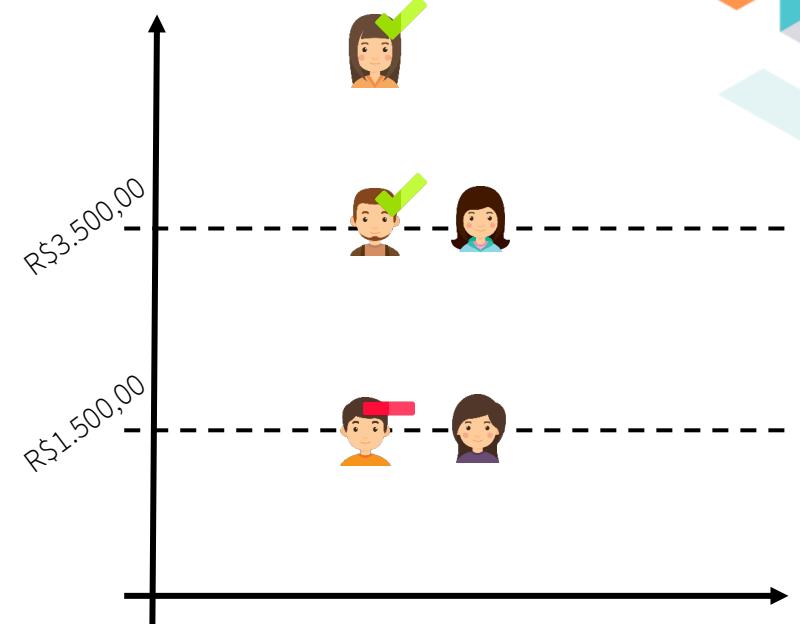
	Nome	Renda	Investidor?	
	João	R\$3.500,00	Sim	
	Luiz	R\$1.500,00	Não	
	Maria	R\$5.000,00	Sim	

Atributos **Classe**

Instâncias

Inteligência Artificial

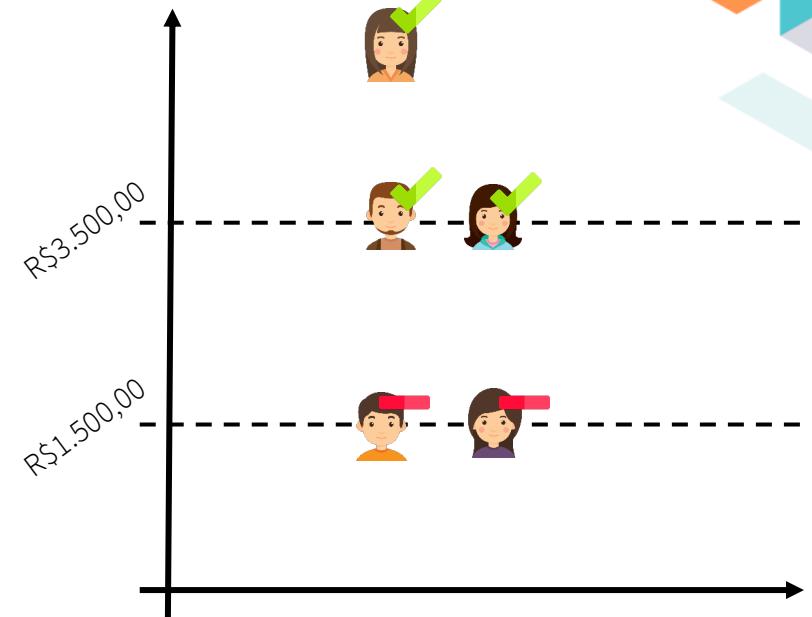
	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Rosa	R\$1.500,00	???
	Ana	R\$3.500,00	???



Novas entradas desconhecidas

Inteligência Artificial

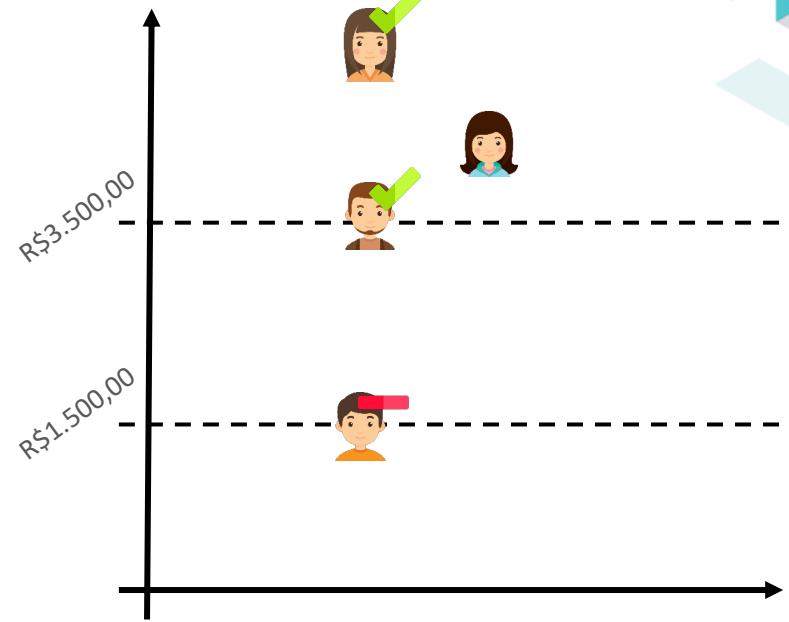
	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Rosa	R\$1.500,00	Não 
	Ana	R\$3.500,00	Sim 



É possível classificar.

Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 
	Luiz	R\$1.500,00	Não 
	Maria	R\$5.000,00	Sim 
	Ana	R\$3.659,00	???



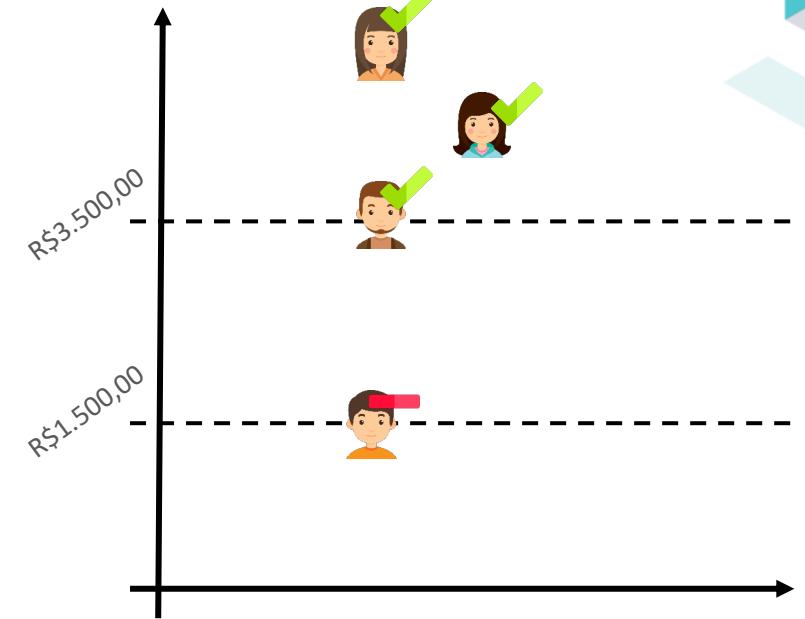
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
-----------------------------------------------------------------------------------	------	-------------	-----------------------------------------------------------------------------------------

	Maria	R\$5.000,00	Sim 
-----------------------------------------------------------------------------------	-------	-------------	-----------------------------------------------------------------------------------------

	Ana	R\$3.659,00	Sim 
-------------------------------------------------------------------------------------	-----	-------------	-------------------------------------------------------------------------------------------



É possível classificar.

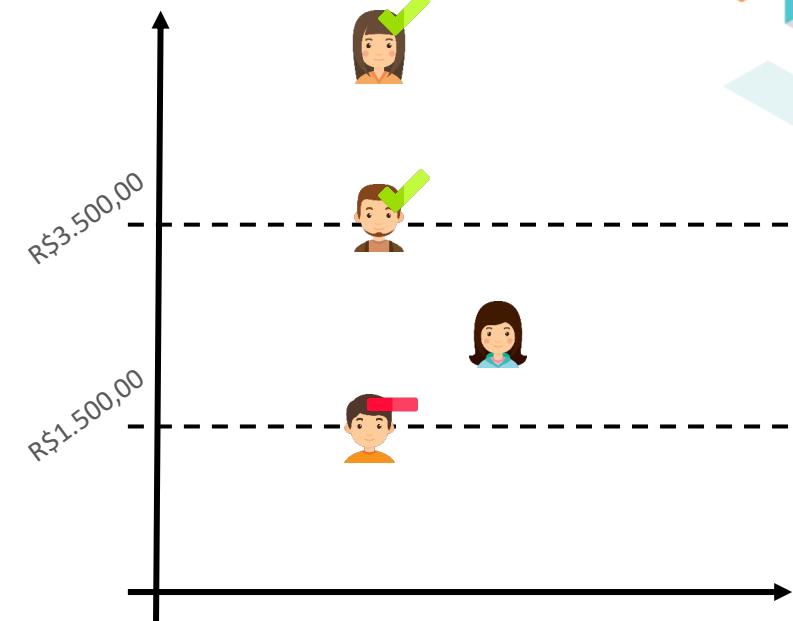
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
-----------------------------------------------------------------------------------	------	-------------	-----------------------------------------------------------------------------------------

	Maria	R\$5.000,00	Sim 
-----------------------------------------------------------------------------------	-------	-------------	-----------------------------------------------------------------------------------------

	Ana	R\$2.000,00	???
-------------------------------------------------------------------------------------	-----	-------------	-----



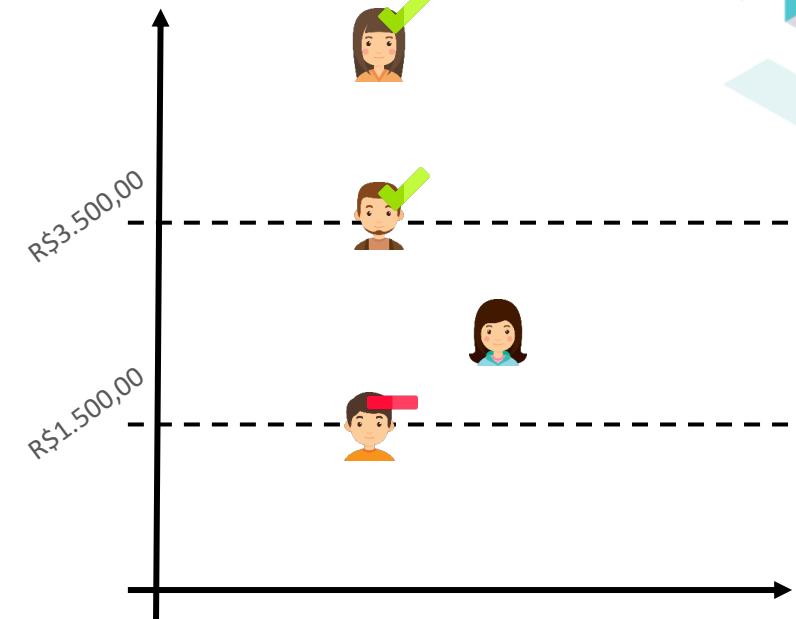
Inteligência Artificial

	Nome	Renda	Investidor?
	João	R\$3.500,00	Sim 

	Luiz	R\$1.500,00	Não 
-----------------------------------------------------------------------------------	------	-------------	-----------------------------------------------------------------------------------------

	Maria	R\$5.000,00	Sim 
-----------------------------------------------------------------------------------	-------	-------------	-----------------------------------------------------------------------------------------

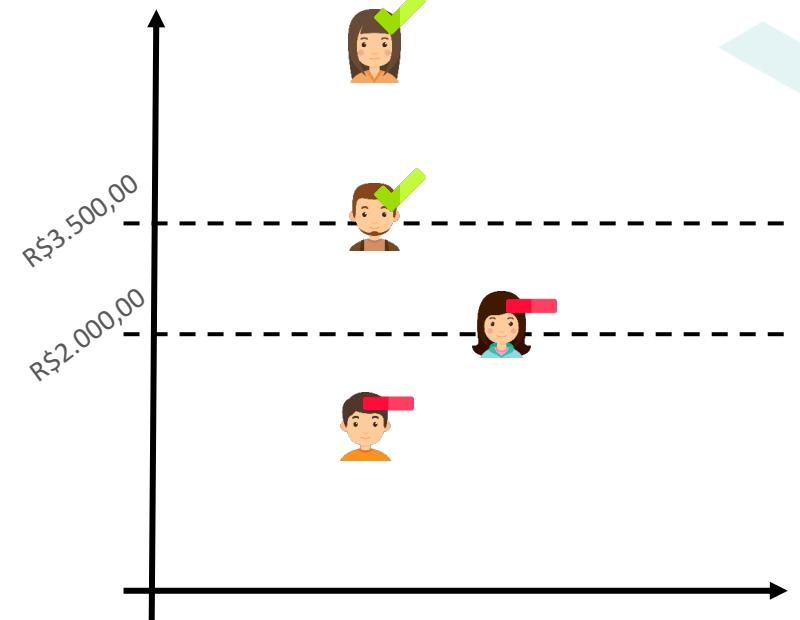
	Ana	R\$2.000,00	???
-------------------------------------------------------------------------------------	-----	-------------	-----



**Não é possível
classificar**

Inteligência Artificial

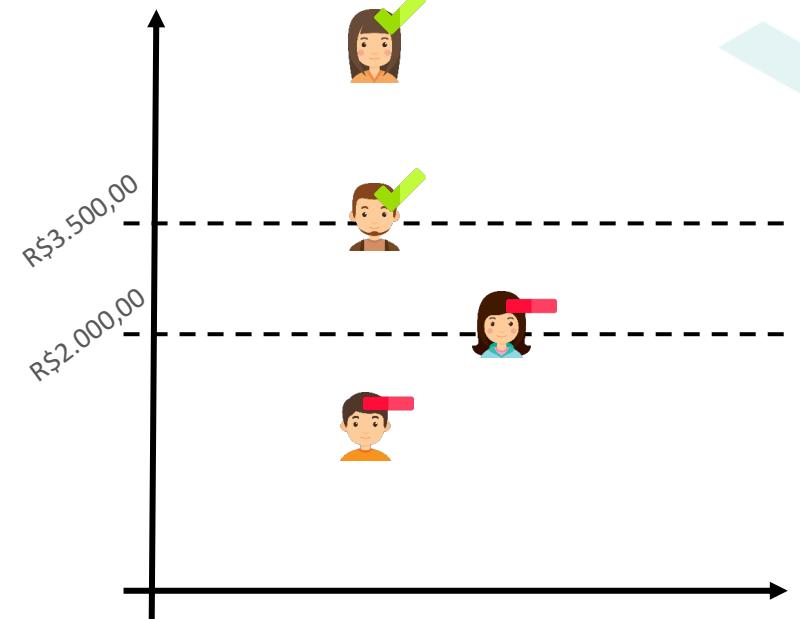
Nome	Renda	Investidor?
João	R\$3.500,00	Sim ✓
Luiz	R\$1.500,00	Não -
Maria	R\$5.000,00	Sim ✓
Ana	R\$2.000,00	Não -

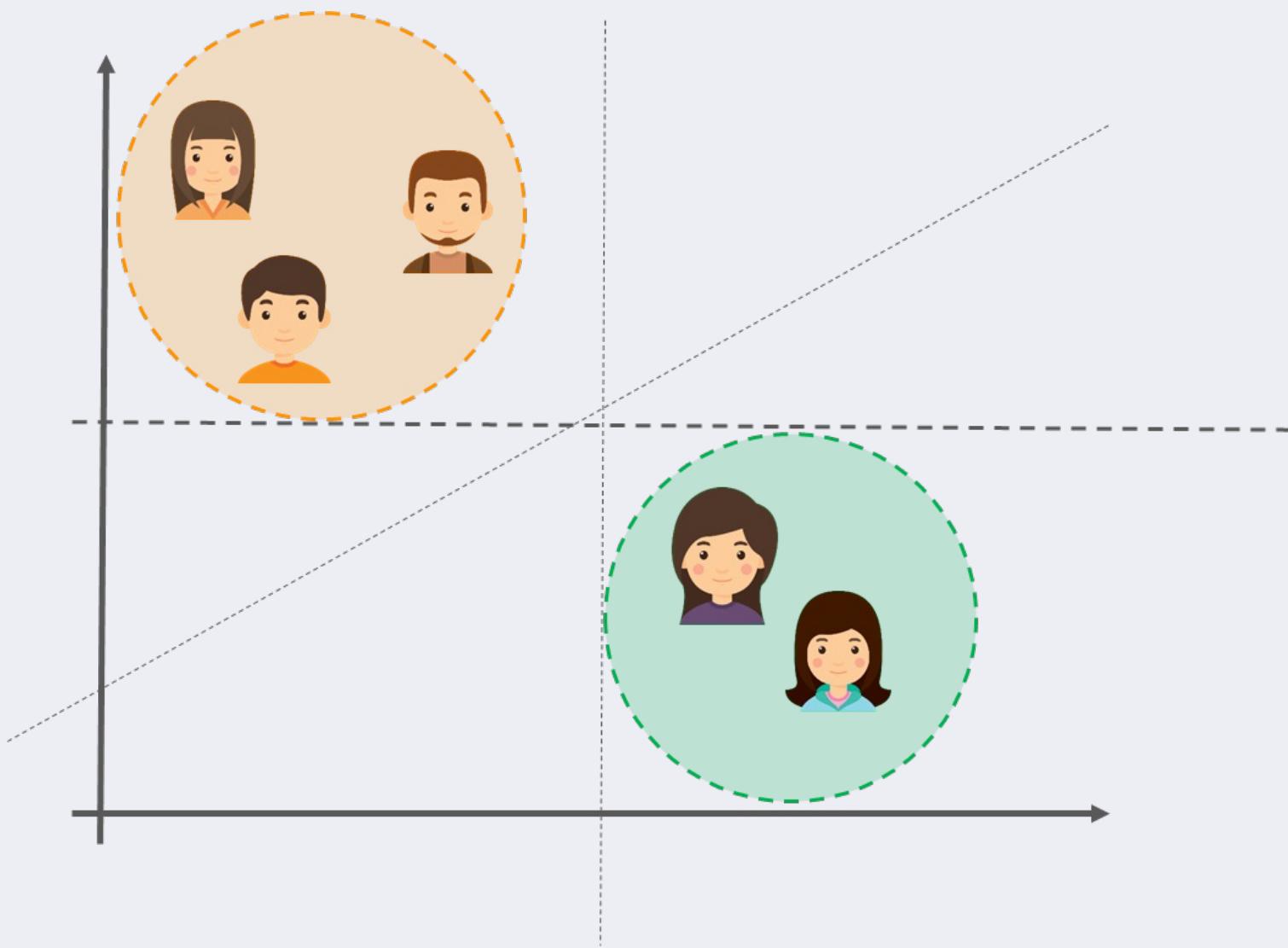


Podemos inserir novos
exemplos

Inteligência Artificial

Nome	Renda	Investidor?
 João	R\$3.500,00	Sim 
 Luiz	R\$1.500,00	Não 
 Maria	R\$5.000,00	Sim 
 Ana	R\$2.000,00	Não 
 José	R\$1.735,00	???





Afinal, como funciona a Inteligência Artificial?



Nome	Renda	Patrimônio	Residência	Já teve investimento	Cliente a quanto tempo	Escolaridade	...	Investidor?
 João	R\$4.500,00	R\$ 1 M	VI. Mariana	Sim	2 anos	Médio	...	Não 
 Luiz	R\$2.500,00	R\$ 250 Mil	Jabaquara	Não	5 anos	Superior	...	Não 
 Maria	R\$6.000,00	R\$ 1.5 M	Morumbi	Não	6 anos	Médio	...	Sim 
 Ana	R\$7.000,00	R\$ 800 Mil	Jardins	Sim	3 anos	Superior	...	Sim 
 José	R\$5.000,00	R\$ 1.2 M	Conceição	Não	2 anos	Superior	...	???

Afinal, como funciona a Inteligência Artificial?

Mas os problemas do mundo real são mais complexos...







NÃO PRECISA CODIFICAR ELE!



What about a dog and a mop? Easy, right?



Not so fast...



Entender não é tão simples

Google me mostre fotos de raças cachorro exceto beagles

Tudo Imagens Compras Vídeos Notícias Mais Definições Ferramentas

filhote raças caninas

Beagle (Beagle Inglês ou famoso Snoopy...) cachorrogato.com.br

Beagle – Raças Caninas - Raças de Cac... petvale.com.br

476 x 238
Beagle (Beagle Inglês ou famoso Snoopy) - Raças de... cachorrogato.com.br

Beagle | Raças de cães | Royal Canin royalcanin.pt

Beagle - Blog do Cachorro blogdocachorro.com.br

Nós, os Cachorros - N... nososcachorros.blogspot...

Beagle: O cachorro cantor | Au au au!!! caninablog.wordpress.com

Beagle: tudo sobre a raça em um guia co... arbolez.com

Raças: Beagle | BitCão bitcao.com.br

Entender não é tão simples

***“Eu vi um homem na
montanha com um telescópio”***



1

*Eu vi um homem. O homem estava na
montanha. Eu estava com o telescópio.*

2

*Eu vi um homem. Eu estava na
montanha. O homem estava com o
telescópio.*

3

*Eu vi um homem. O homem estava na
montanha. O homem estava com o
telescópio.*

4

*Eu vi um homem. Eu estava na
montanha. Eu estava com o telescópio.*

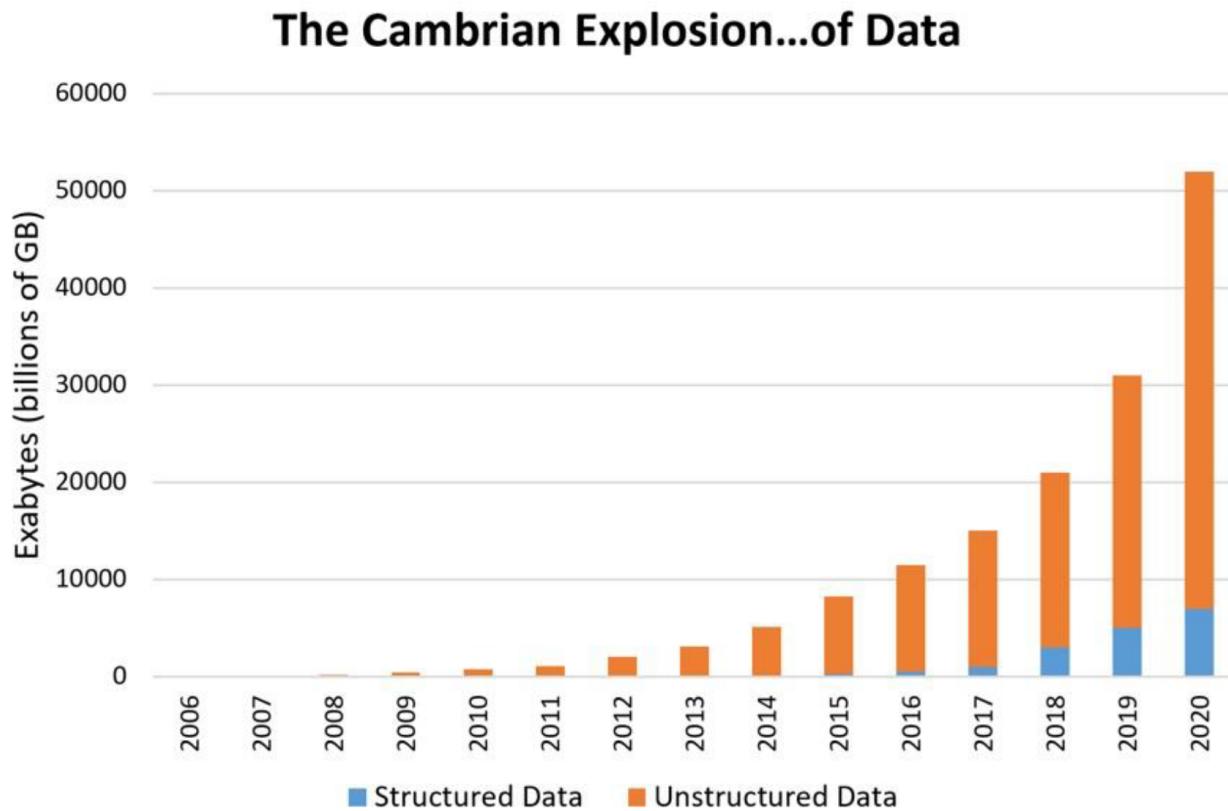


Figura 16 – Taxa de expansão na produção de dados estruturados e não-estruturados

Fonte: https://www.eetimes.com/author.asp?section_id=36&doc_id=1330462

- 
- 1 Motivação e **Contextualização Geral**
 - 2 Ciclo de vida do **dado**
 - 3 Ciclo de vida do **modelo**
 - 4 Técnicas e **Algoritmos**
 - 5 Perfil e responsabilidades dos **profissionais de dados**
 - 6 Aplicação e **Cases**



The world's
most valuable
resource is no
longer oil,
but data

Produtos que possuem petróleo como matéria-prima



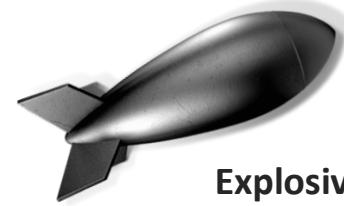
Óleos Lubrificantes para motores



Pneus e borrachas



**Diesel e
Gasolina**



Explosivos



Piche para asfalto



Tecidos de poliéster e poliamida



Collage



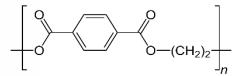
Tintas



Materiais Plásticos



Garrafas PET



Uma verdadeira infinidade de produtos

Por que **DADOS** são
importantes?



Organização orientada para o consumo de dados

no passado

Empresas criam produto e pessoas consomem o que lhes é oferecido

Valor = conversão de matéria prima em produtos



Crescimento:
Aumento de preço ou quantidade produzida



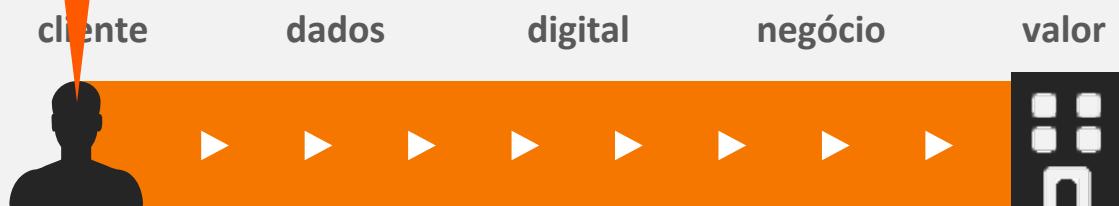
Dados usados no suporte para tomada de decisão
Receita define a perenidade do negócio

hoje

Dados revelam novas demandas de negócio que a tecnologia ajuda a materializar



Insights:
Dados revelam as demandas dos clientes



Novo negócio para transformar mercado
Relevância para o cliente que define a perenidade do negócio

Como organizar os **DADOS?**



Conceito de fluxo de extração de petróleo



-  1
-  2
-  3
-  4
-  5
-  6

Ciclo de vida do Dado

Conceito de fluxo de extração de petróleo

1 CRIAÇÃO

DADO TRANSACIONAL

- Criação do dado
- Conceito e metadados



Ciclo de vida do Dado

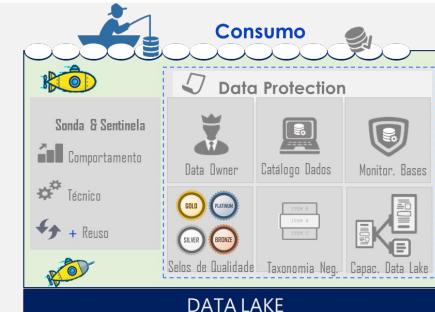
Conceito de fluxo de extração de petróleo

2 CAPTURA E AQUISIÇÃO

- Democratização do dado
- Definição de confidencialidade e privacidade
- Atribuição de owner
- Controle, monitoramento e indicadores

RAW DATA

Source of Record (SOR)



DATALAKE – CAMADA RAW (SOR)

A	14/06/2019	10:00:05	161789046785903	0010098700
B	14/06/2019	11:50:99	759999876990900	0000900000
A	14/06/2019	12:30:05	313356889567788	0000015900

Campo	Tipo	Descrição
TIP_TRAN	Char(1)	Tipo transação Domínio: A - SAQUE COM CARTAO MAGNETICO EM CONTA B - DEPOSITO DINHEIRO EM C/C-C/P AUTOATEND.
DAT_TRAN	Char(8)	Data transação
HOR_TRAN	Char(8)	Hora transação
NUM_ID_TRAN	Num(15)	Núm. Identificação da transação
VAL_ORGL_TRAN	Decimal(10)	Valor original da transação

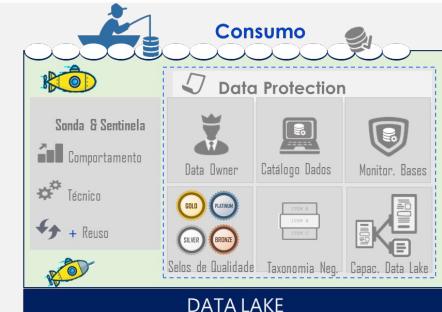
Ciclo de vida do Dado

Conceito de fluxo de extração de petróleo

3 TRANSFORMAÇÃO DOS DADOS

- Harmonização
- Padronização
- Higienização dos dados
- Equalizador de formato
- Consolidador de dados

Source of Truth (SOT)



DATALAKE - SOT

01	14/06/2019	10:00:05	161789046785903	100.987,00
02	14/06/2019	11:50:99	759999876990900	9.000,00
01	14/06/2019	12:30:05	313356889567788	159,00

BASE ORIGEM

Campo	Tipo	Domínio
Tipo de transação	Char(1)	A - SAQUE COM CARTAO MAGNETICO EM CONTA B - DEPOSITO DINHEIRO EM C/C-C/P AUTOATEND.

DOMÍNIO CORPORATIVO

Campo	Tipo	Descrição
Tipo de transação	Num(2)	01 - SAQUE COM CARTAO MAGNETICO EM CONTA 02 - DEPOSITO DINHEIRO EM C/C-C/P AUTOATEND.

Conceito de fluxo de extração de petróleo



4 ENRIQUECIMENTO DOS DADOS

- Controle de qualidade (visão negócio)
- Agregação de dados
- Criação de variáveis
- Aplicação de regra de negócios

DADO
ESPECIALIZADO

BOOK DE VARIÁVEIS



Conceito de fluxo de extração de petróleo



5 GERAÇÃO DE INSIGHTS

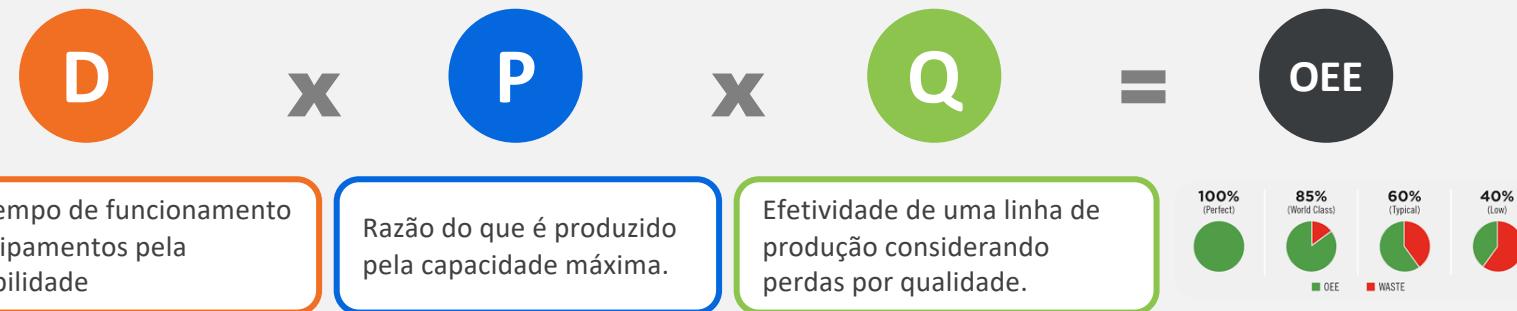
- Criação de:
 - Analise exploratórias
 - Indicadores
 - Dashboards
 - Modelos
 - Algoritmos
 - Machine Learning

ANALYTICS e
ADVANCED ANALYTICS



Exemplo: OEE – Definição do Indicador

Overall equipment effectiveness (OEE) é um indicador que quantifica o desempenho de uma unidade de fabricação em relação à sua capacidade projetada, durante os períodos em que ela está programada para ser executada.

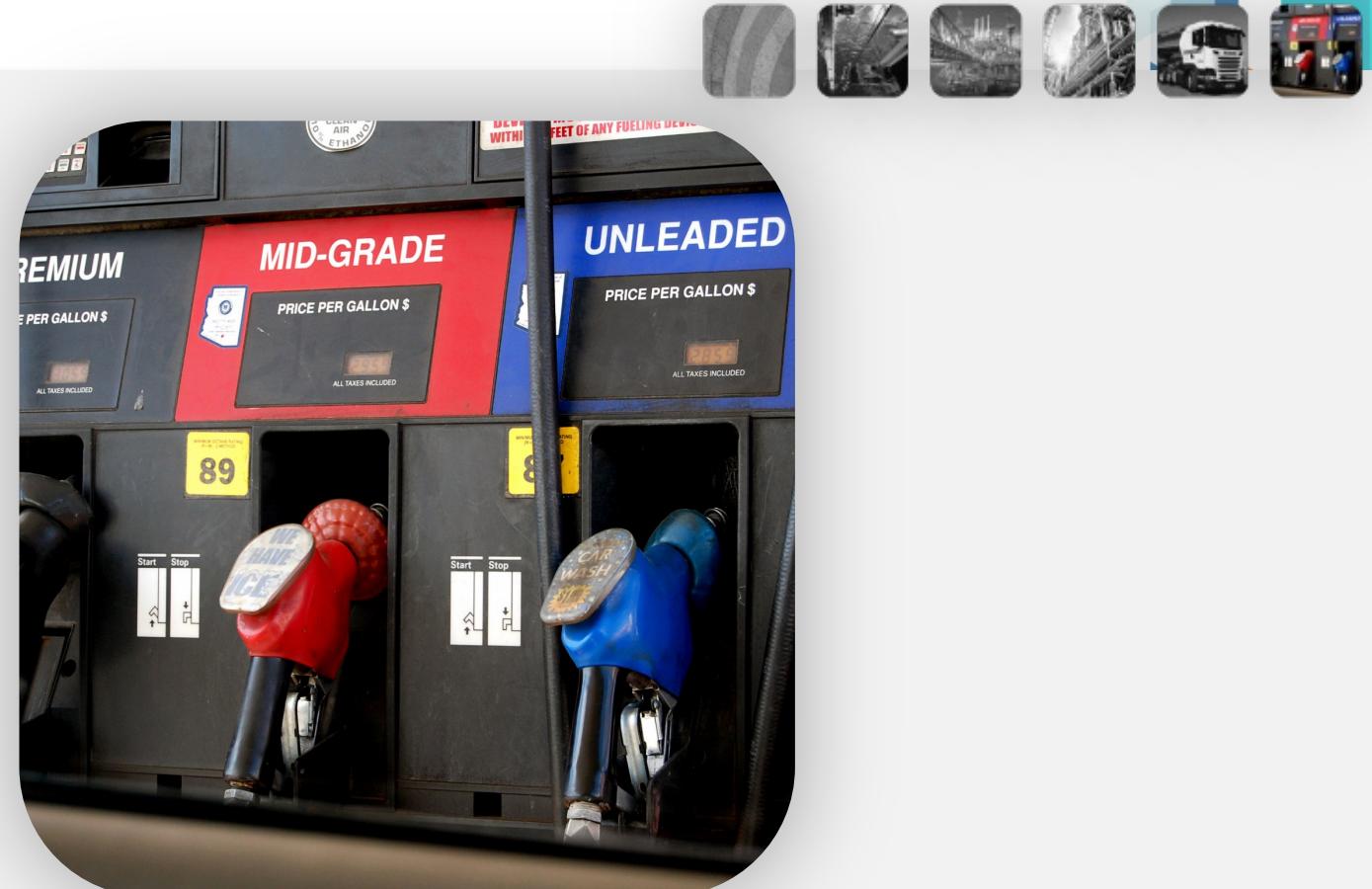


Conceito de fluxo de extração de petróleo

6 TOMADA DE DECISÃO COM OS DADOS

- Decisões baseadas em dados

Ciclo de vida do dado se completa no uso para tomada de decisão



- 
- 1 Motivação e **Contextualização Geral**
 - 2 Ciclo de vida do **dado**
 - 3 Ciclo de vida do **modelo**
 - 4 Técnicas e **Algoritmos**
 - 5 Perfil e responsabilidades dos **profissionais de dados**
 - 6 Aplicação e **Cases**

Qual o ciclo de vida de um modelo?

Ciclo de vida de Modelos em Machine Learning



#1 – Análise e estudo do problema

Pipeline de projetos

- Análises dos dados
- Construção das variáveis

Bases estruturadas
dados não estruturados



Variável	Nome	Variável resposta
1	Pablo	A
2	Vinicio	B
3	Alex	B
4	Rogers	A
5	Pedro	A

#2 – Extração e captura do dado

Pipeline de projetos

- Amostragem e transformação
- Medidas de poder preditivo

Variável	Nome	Variável resposta
1	Pablo	A
2	Vinicius	B
3	Alex	B
4	Rogers	A
5	João	B
6	Antonio	A
7	Pedro	A

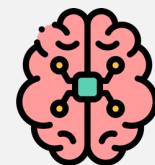
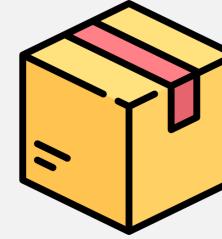
Variável	Nome	Variável resposta
1	001	A
2	010	B
3	011	B



#3 – Refino e definição do escopo

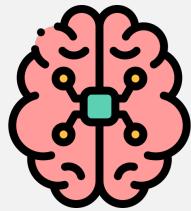
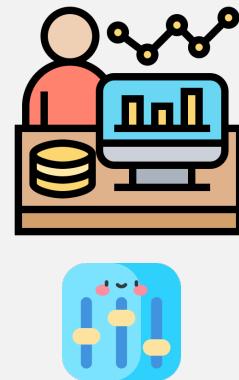
- Delimitação do escopo
- Testes estatísticos
- Validação do modelo

Variável	Nome	Variável resposta
1	001	A
2	010	B
3	011	B



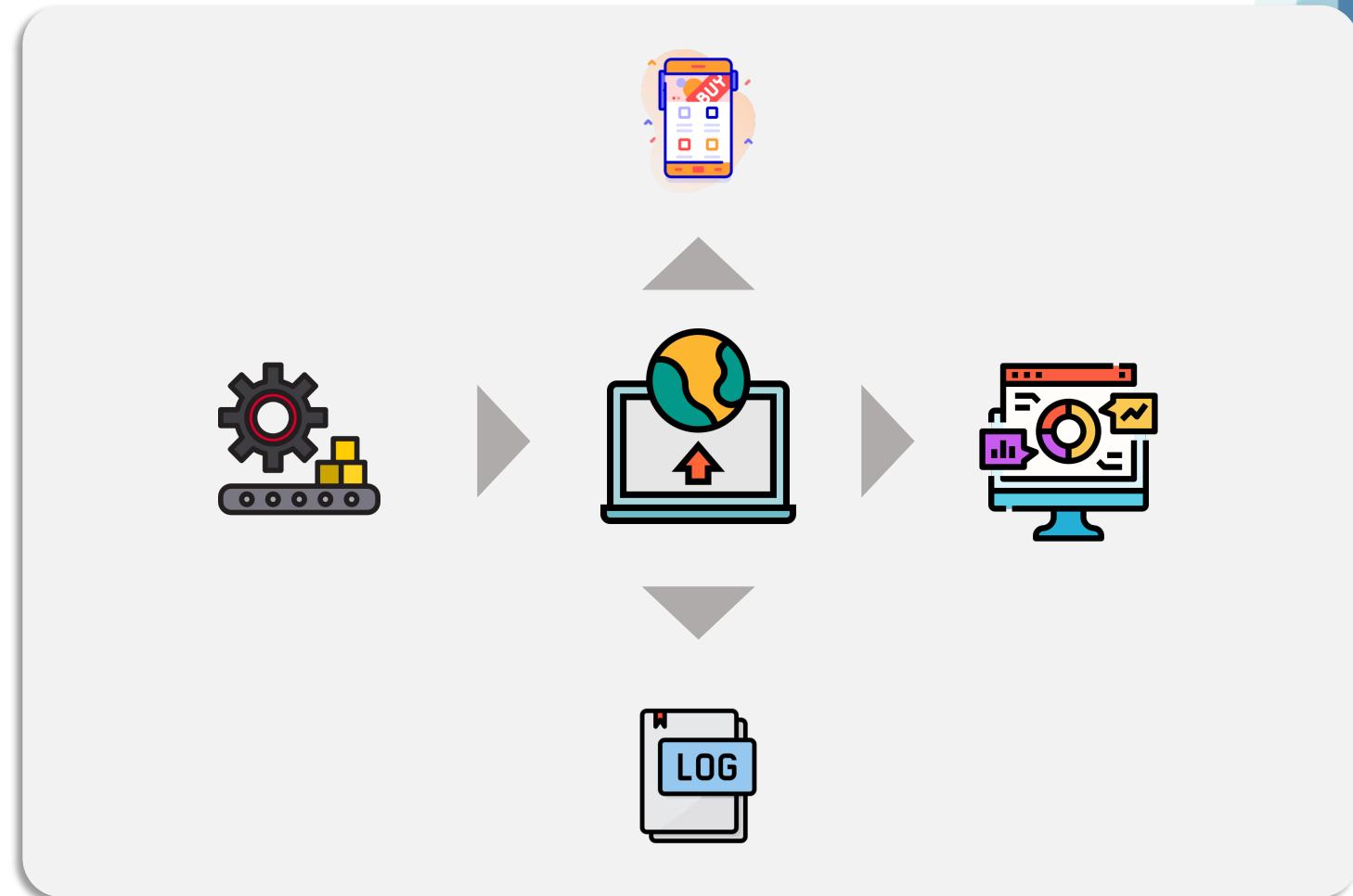
#4 – Gemcrafting

- Calibração de variáveis
- Calibração de parâmetros
- Validação das hipóteses
- Redesenho



#5 – Deploy

- Montagem da infraestrutura
- Adequação a esteira de deploy
- Métricas de tempo de resposta
- Métricas de acompanhamento



Como fazer um pão modelo?



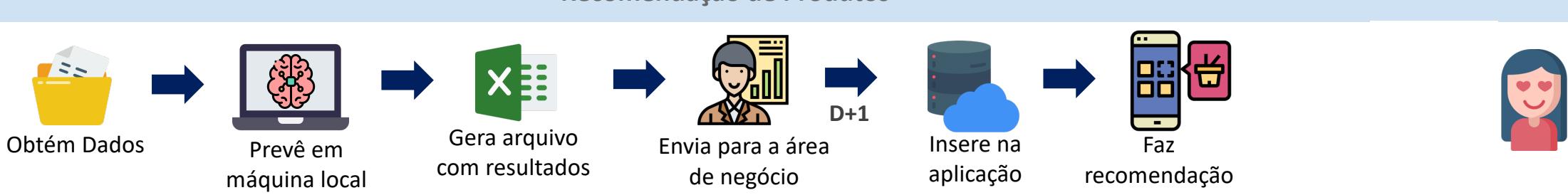


“More than 87% of data science projects never make it into production”

- Multiple studies and surveys -

Modelos em Produção

Batch

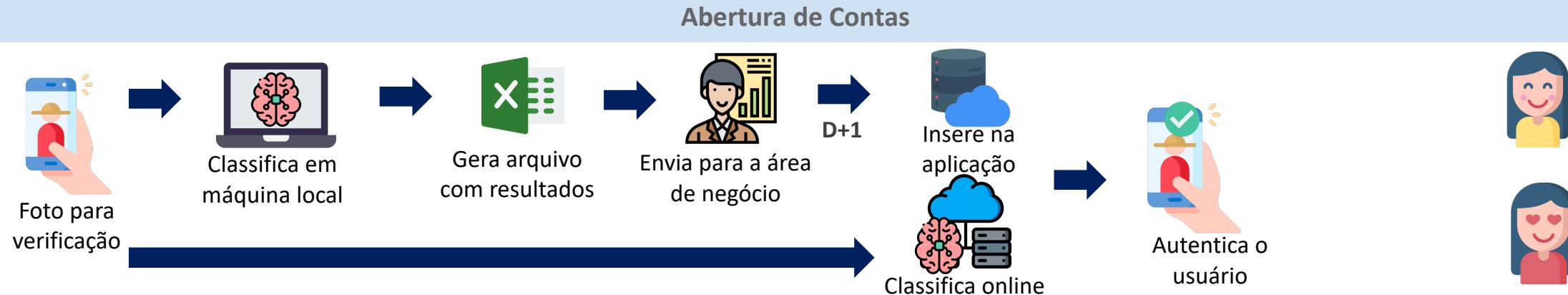


Modelos em Produção

Batch

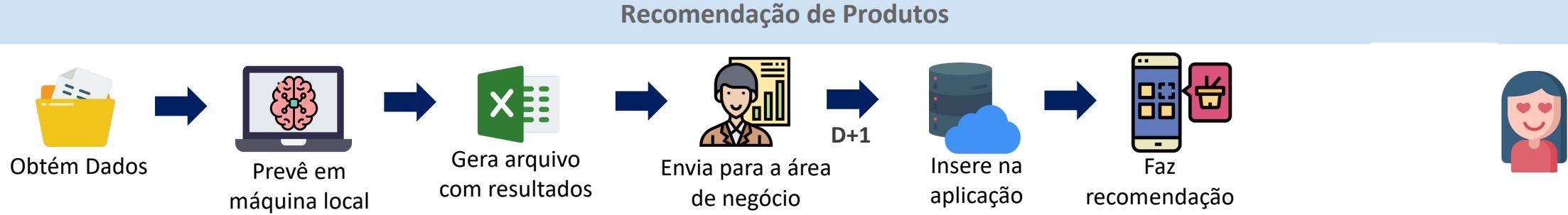


Batch ou Online

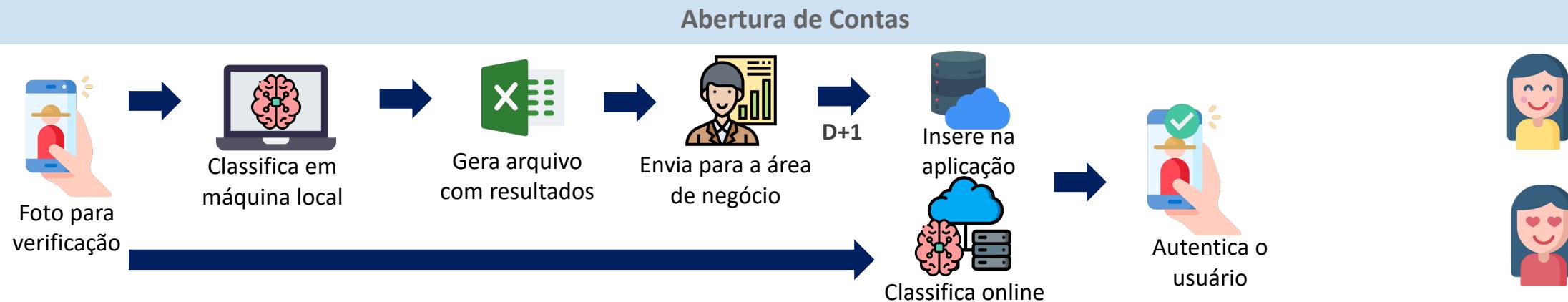


Modelos em Produção

Batch



Batch ou Online



Online



If ML is a rocket engine,
data is the fuel



**Launching is easy,
Operating is hard.**

**"The real problems with a
ML system will be found
while you are continuously
operating it for the long term"**



“

Simples ML + More data
> Fancy ML + Small data

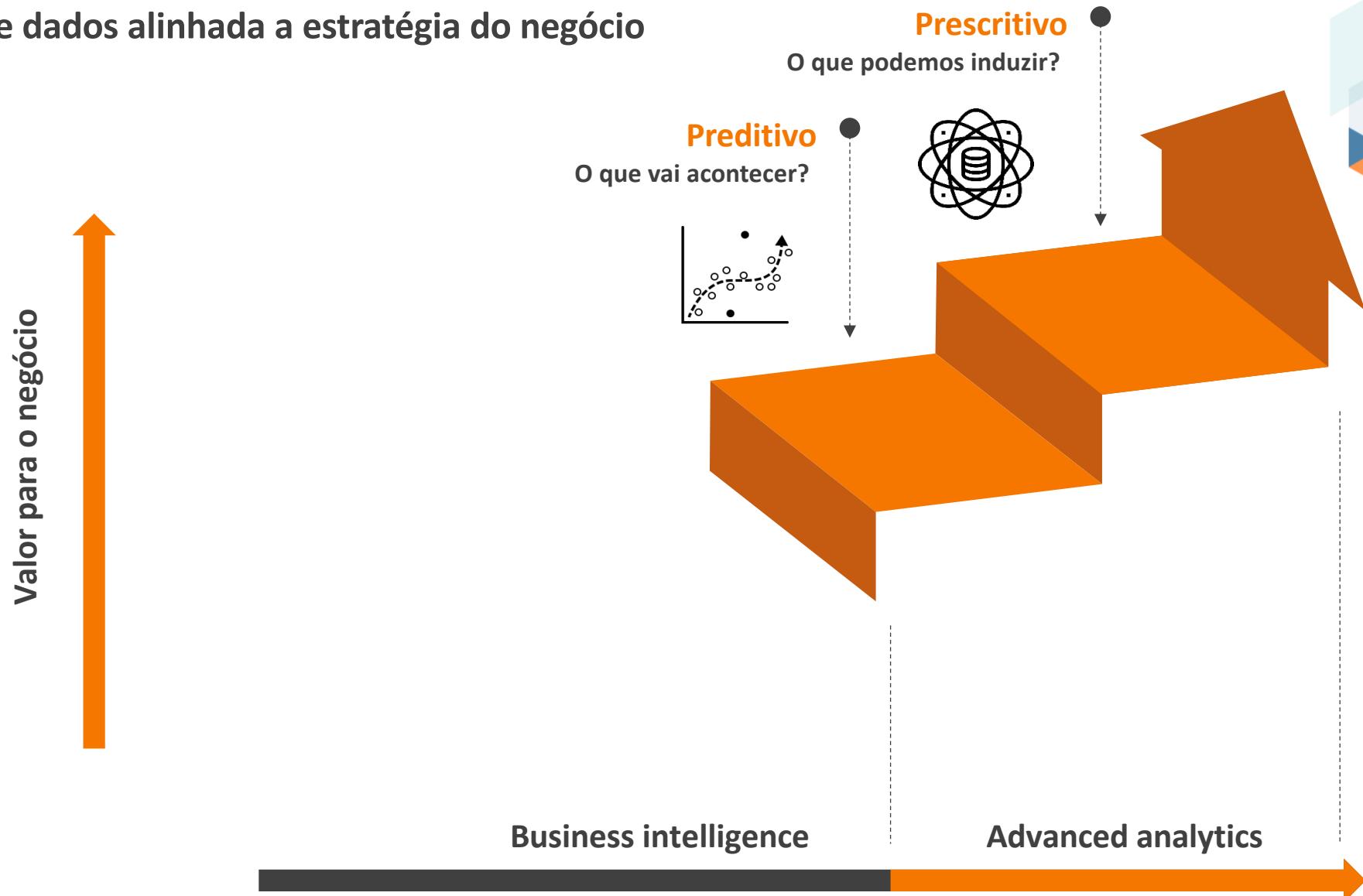
- 
- 1 Motivação e **Contextualização Geral**
 - 2 Ciclo de vida do **dado**
 - 3 Ciclo de vida do **modelo**
 - 4 Técnicas e **Algoritmos**
 - 5 Perfil e responsabilidades dos **profissionais de dados**
 - 6 Aplicação e **Cases**

E como tem evoluído o uso de
ANALYTICS e geração de valor
para os negócios?



Evolução do uso de Analytics

Estratégia de dados alinhada a estratégia do negócio



Evolução do uso de Analytics

Estratégia de dados alinhada a estratégia do negócio





What is descriptive analytics?



Descriptive Analytics

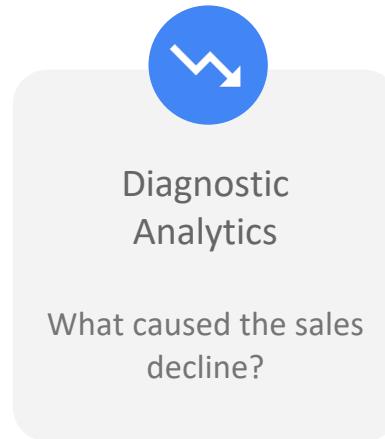
How is the business performing?

Describes state of the business

Helps to answer questions such as...

- ▶ *How is our business performing?*
- ▶ *Who are our best customers?*
- ▶ *What are our best selling products?*

What is diagnostic analytics?



Analysis that helps to diagnose issues and root cause

Used to answer questions such as...

- ▶ *What caused a decline in sales?*
- ▶ *Why did a region miss its target?*

What is predictive analytics?



Predictive Analytics

What is the sales forecast for next quarter?

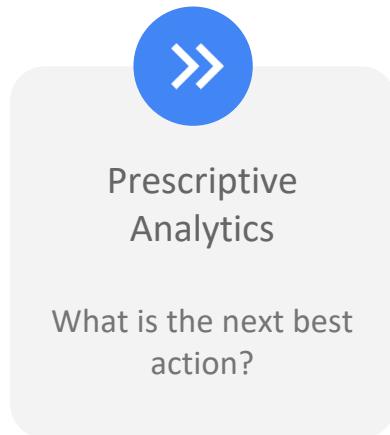
Forward-looking analysis to anticipate the future

Helps to answer questions such as...

- ▶ *What is our sales forecast for next quarter?*
- ▶ *Which customers are likely to default?*
- ▶ *Which prospects are most likely to buy our product?*



What is prescriptive analytics?



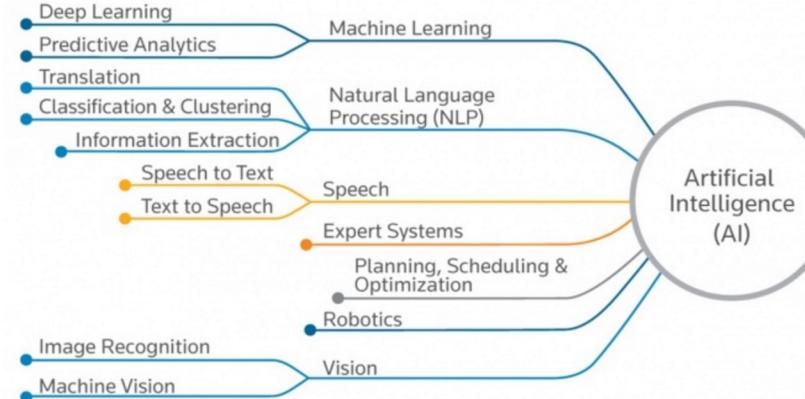
Gives clear recommendation on the best course of action

Helps to answer questions such as...

- ▶ *How should we invest our money?*
- ▶ *What is the next best move? (AlphaGo)*
- ▶ *What is the best route to my destination?*

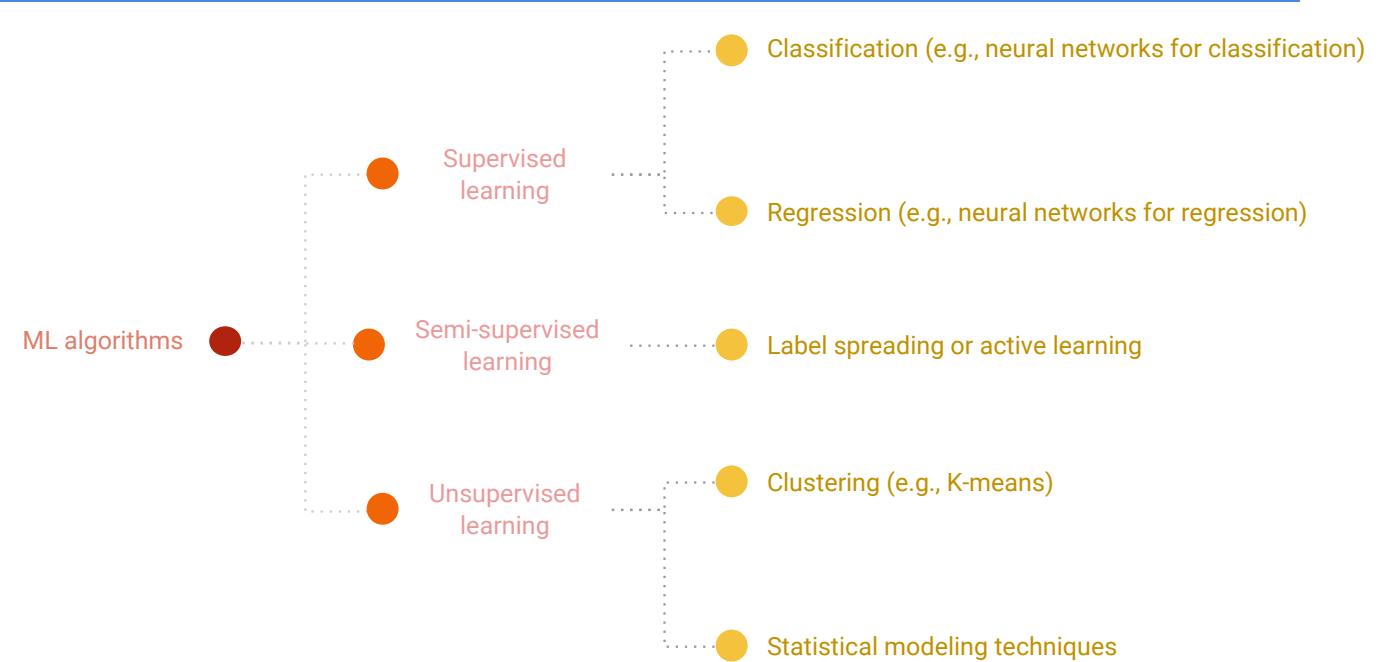
Machine learning is...

-  a branch of artificial intelligence
-  a way to solve problems without explicitly codifying the solution
-  a way to build systems that improve themselves over time



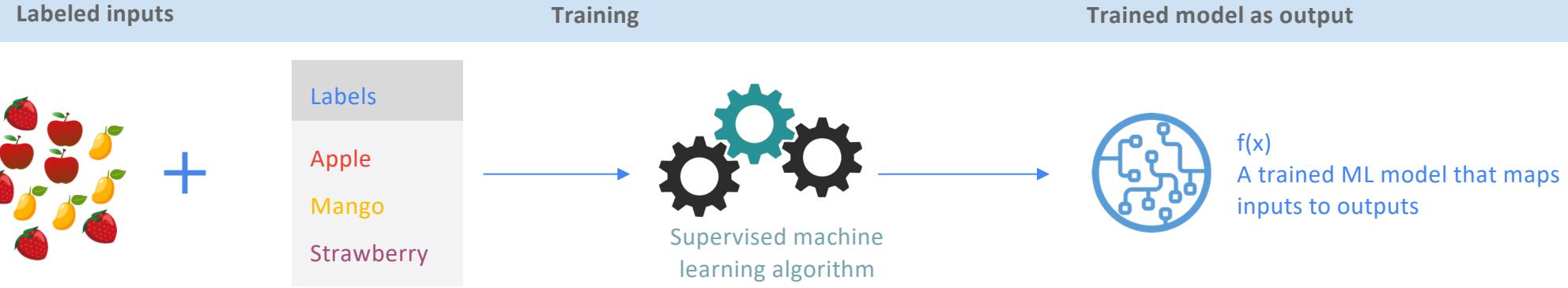
Source: Neota Logic

ML strategy

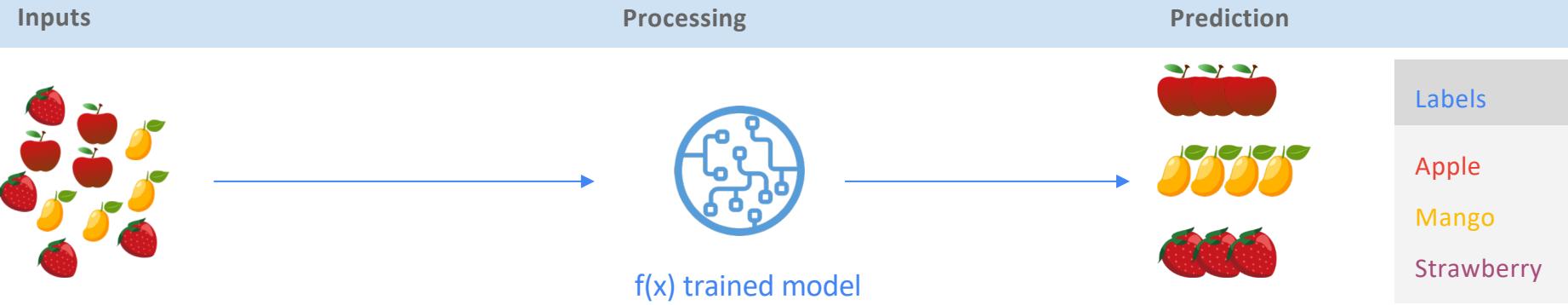


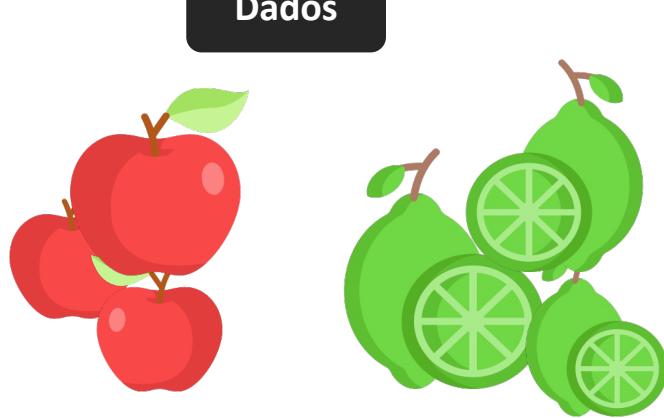
Types of machine learning algorithms | Supervised learning

Training

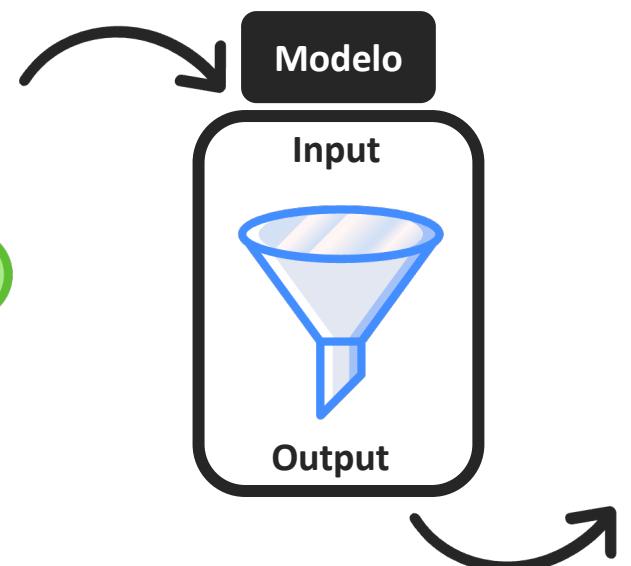


Prediction





Dados



Maçãs



Limões

Types of machine learning algorithms | Unsupervised learning

Training

Labeled inputs



Training

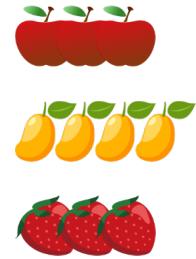


Unsupervised machine learning algorithm

Trained model as output



$f(x)$
Model: Learns clusters or patterns in data to detect anomalies



Prediction

Inputs

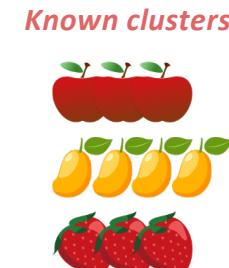


Processing

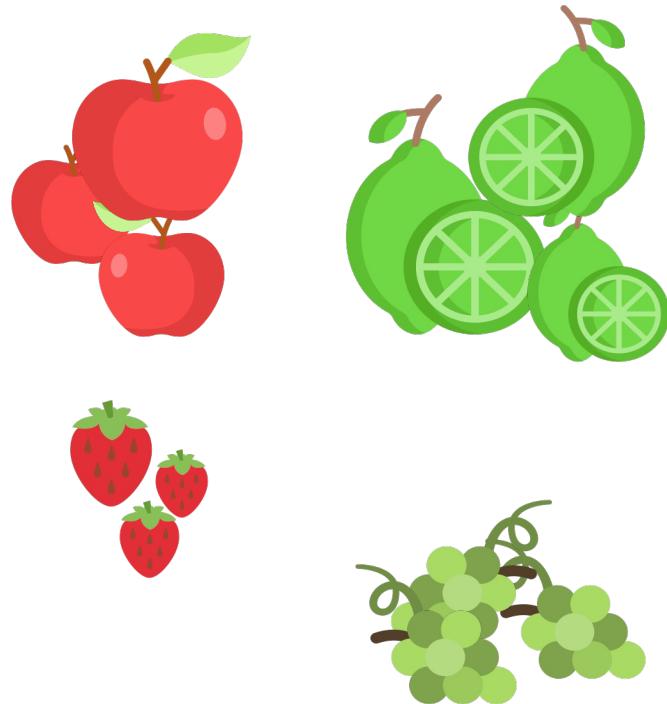


$f(x)$ trained model

Prediction



Dados



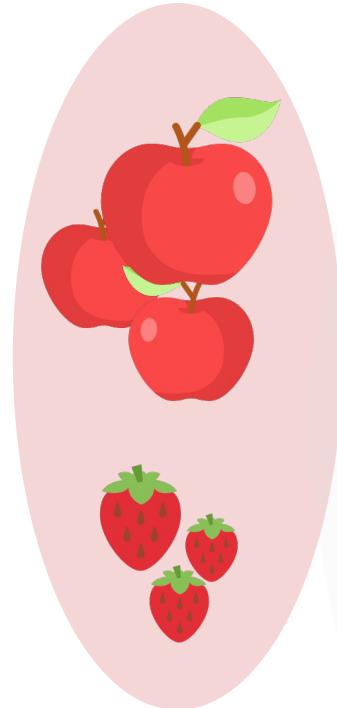
Modelo

Input

Output

Grupo 1

Grupo 2

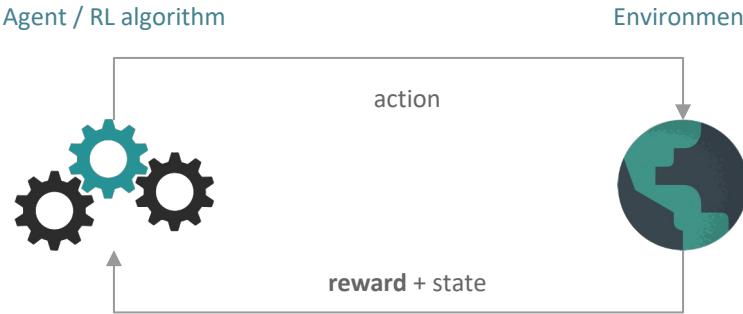


Types of machine learning algorithms | Reinforcement learning

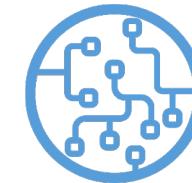
Training

Policy in action

Agent interacting with environment

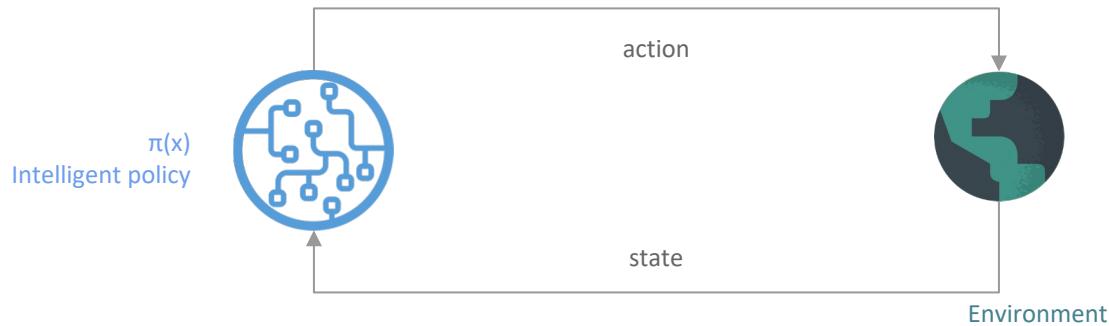


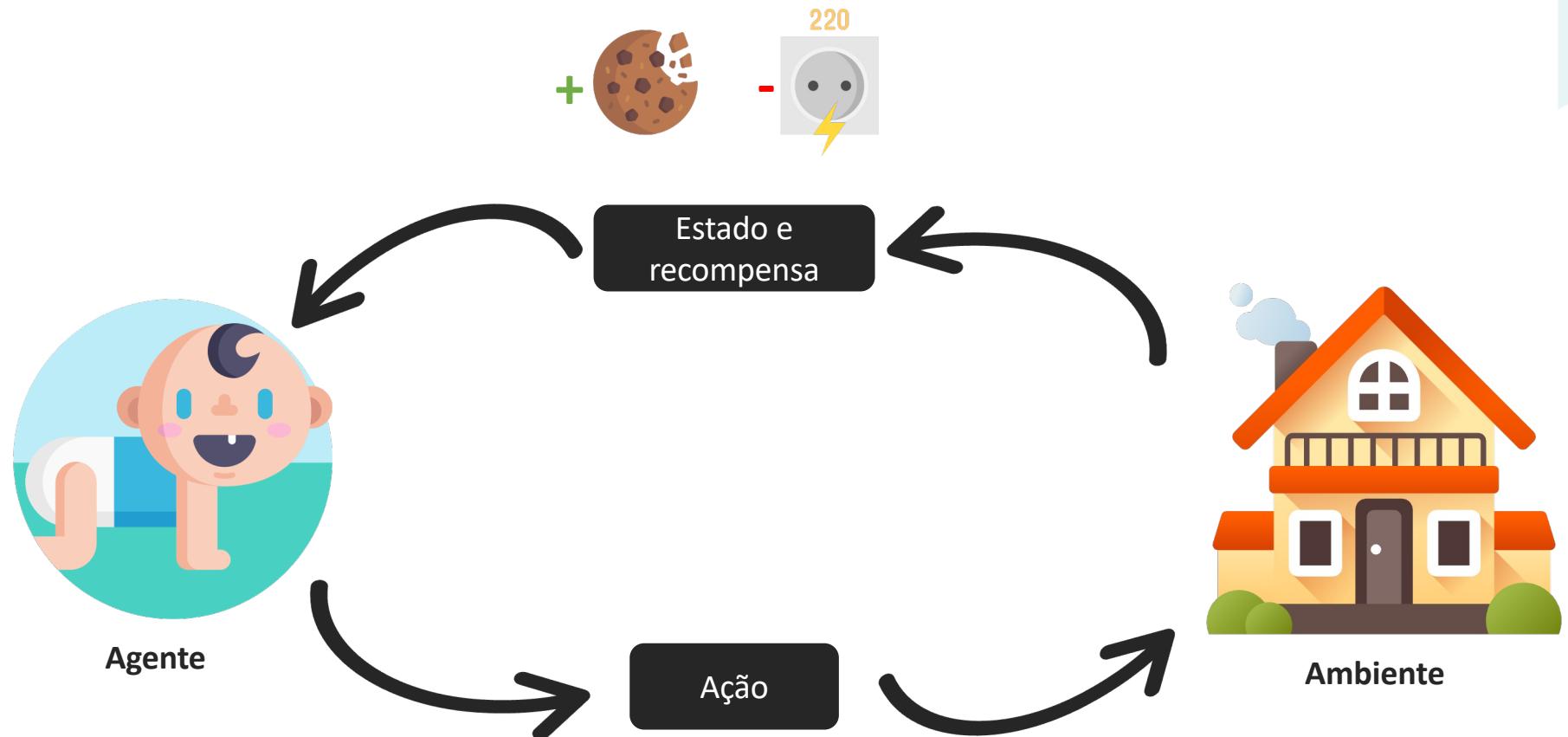
Intelligent Policy



$\pi(x)$
Intelligent policy that understands
environment

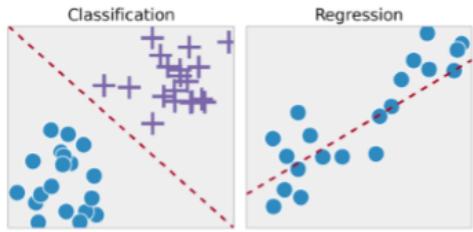
Trained agent / policy intelligently interacting with the environment





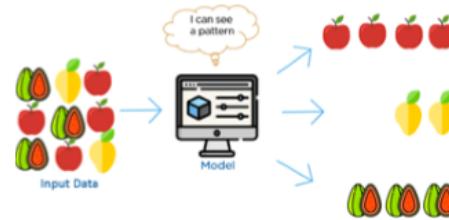
Supervised Learning

- Regression – sales forecasting
- Classification – grouping potential customers



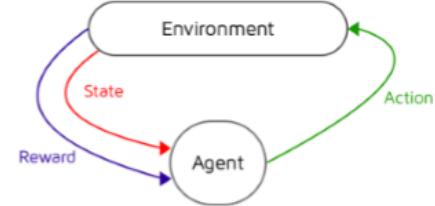
Unsupervised Learning

- Clustering – splitting customers by preferences
- Dimensionality Reduction – storing less data



Reinforcement Learning

- Reasoning - robotics



Deep Learning

- Neural Networks – image recognition



Capacidades

Assistentes pessoais Diagnóstico automatizado Serviço ao cliente Detecção de fraude Produto Smart Pesquisa Automação de Marketing Serviços Analytics Fiscalização Otimização de processo

Visão computacional Reconhecimento de fala Processamento de linguagem natural Sistemas recomendados Máquina de tradução Modelos de aprendizagem personalizados

Machine Learning

Agrupamento Árvores de decisão Algoritmos genéticos Redes Neurais profundas Máquinas de vetores de suporte Regras de associação Redes Bayesianas

- 
- 1** Motivação e **Contextualização Geral**
 - 2** Ciclo de vida do **dado**
 - 3** Ciclo de vida do **modelo**
 - 4** Técnicas e **Algoritmos**
 - 5** Perfil e responsabilidades dos **profissionais de dados**
 - 6** Aplicação e **Cases**

Como conectar usuários para produzir resultados **10x** mais impactantes?



ML Engineer



Data Engineer



Developer



Data Scientist



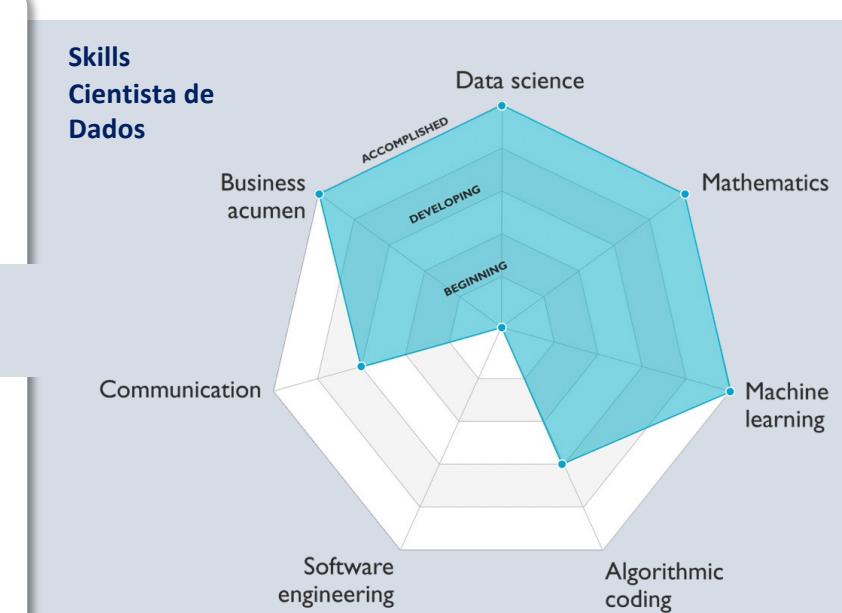
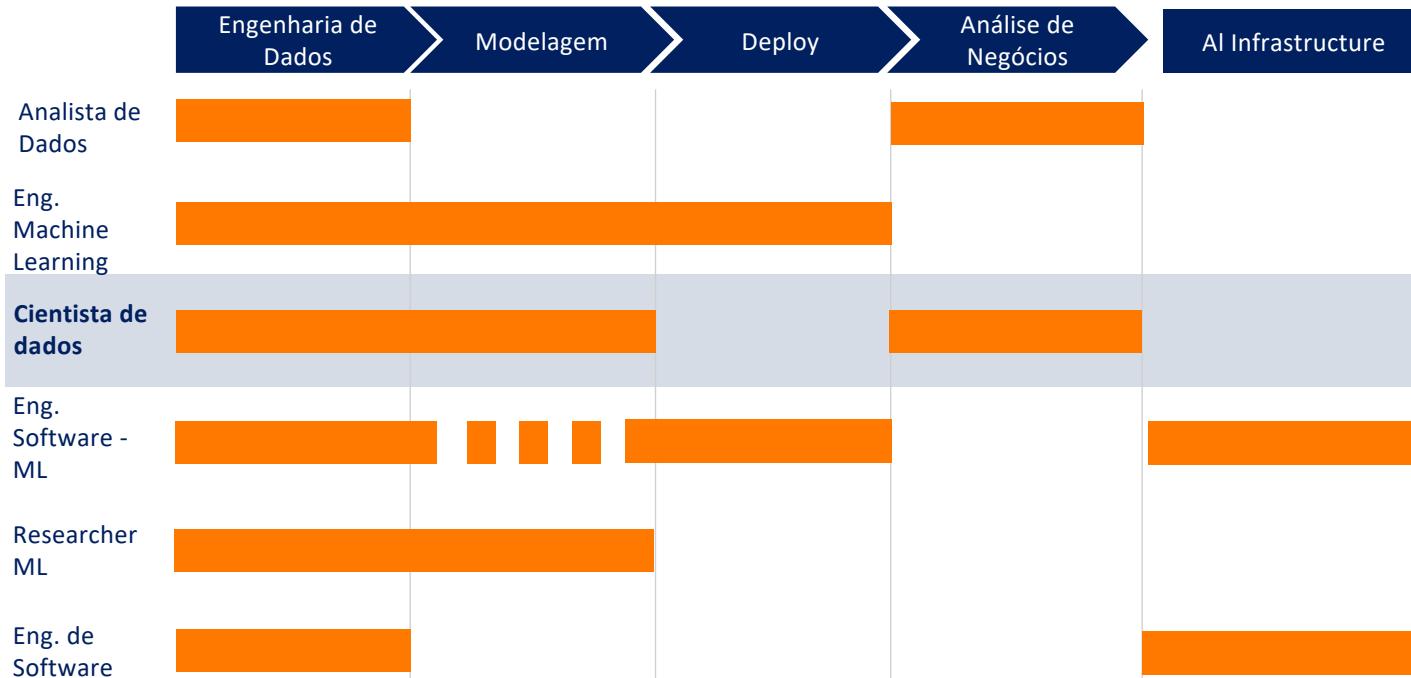
Business Analyst



End User

Perfil do Cientista de dados

Frentes de trabalho x profissionais , de acordo com texto publicado no Workera:



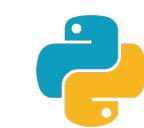
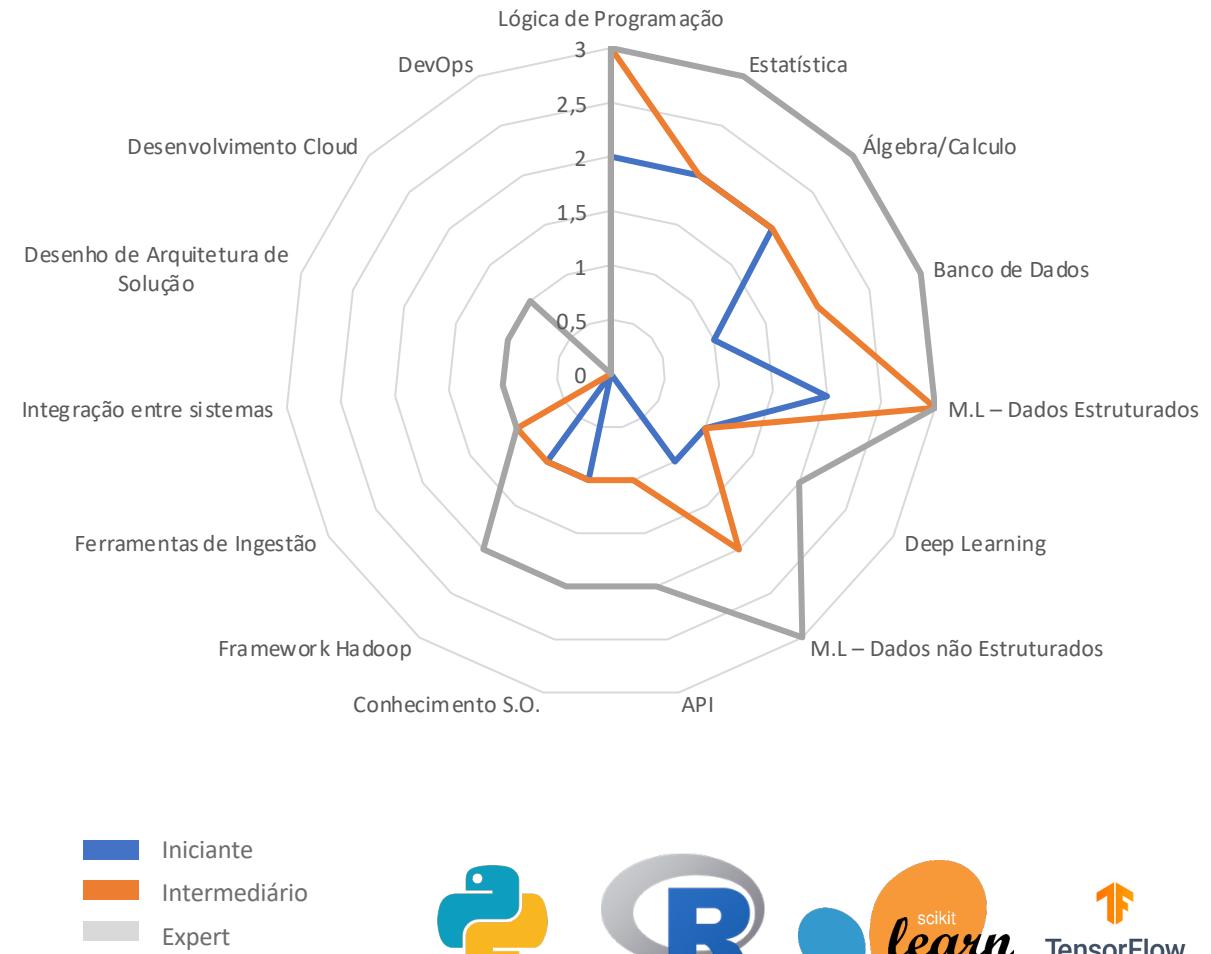
Profissionais x Skills	Data Science	Mathematics	Machine Learning	Algorithmic coding	Software eng.	Communication	Business acumen
Analista de Dados	✓			✓		✓	✓
Eng. Machine Learning	✓	✓	✓	✓	✓		
Eng. Software - ML			✓	✓	✓	✓	
Researcher ML		✓	✓	✓	✓		Research
Eng. de Software				✓	✓	✓	

O Perfil do Cientista de dados

Para cientista de dados é realizado uma prova com duração de 3 horas e nota de corte 5 e após aprovação na prova existe um entrevista técnica (sabatina), e após aprovação em ambas existe a seleção de perfil por vaga (gestor funcional).

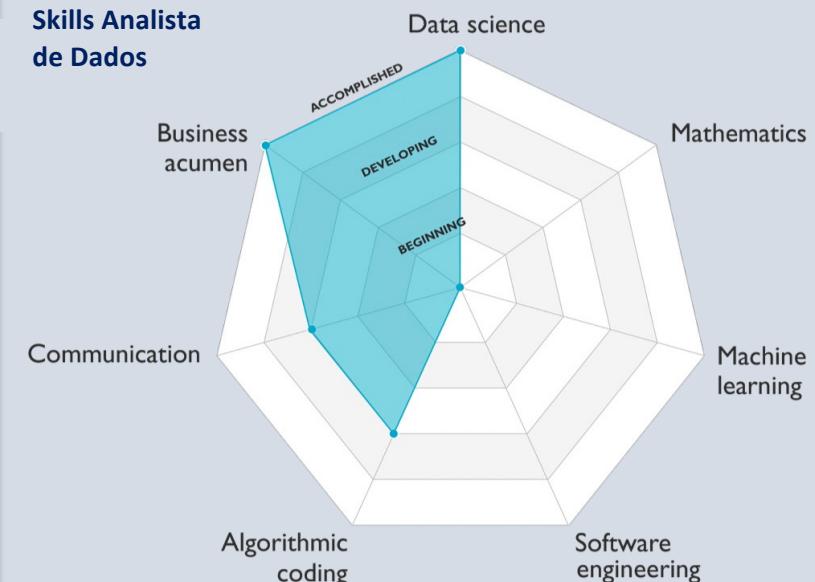
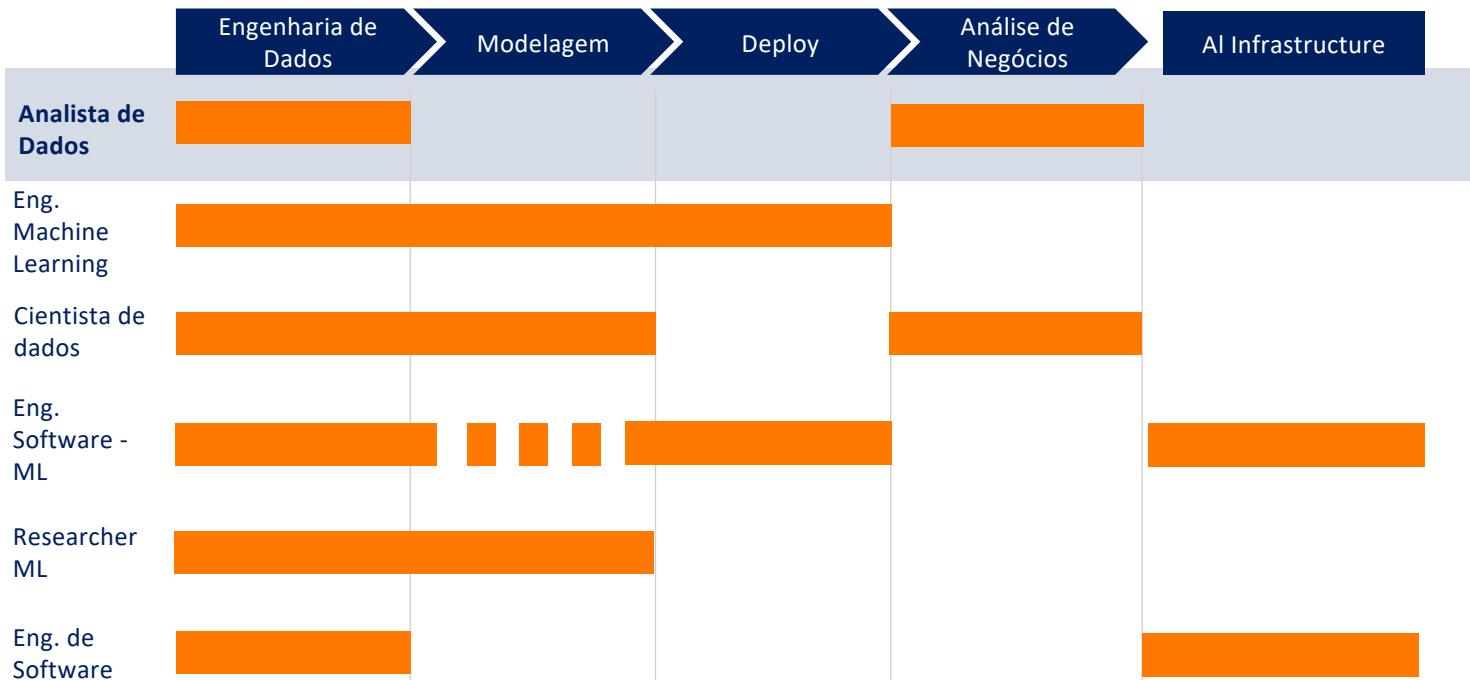
PROCESSOS

Conhecimento	Iniciante	Intermed	Expert
Conhecimento Técnicos			
Lógica de Programação	2	3	3
Estatística	2	2	3
Álgebra/Calculo	2	2	3
Banco de Dados	1	2	3
Machine Learning – Dados Estruturados	2	3	3
Deep Learning	1	1	2
Machine Learning – Dados não Estruturados	1	2	3
Construção e Funcionamento de API	-	1	2
Conhecimento S.O.	1	1	2
Framework Hadoop	1	1	2
Ferramentas de Ingestão	-	1	1
Integração entre sistemas	-	-	1
Desenho de Arquitetura de Solução	-	-	1
Desenvolvimento Cloud	-	-	1
DevOps	-	-	-
Habilidades			
Solução de Problemas	2	2	3
Comunicação	1	2	3
Transformação/Inovação	1	2	2
Visão Estratégica e Sistêmica	1	2	3
Influência	-	-	1
Tutoria	-	2	3



Perfil do Analista de dados

Frentes de trabalho x profissionais , de acordo com texto publicado no Workera:

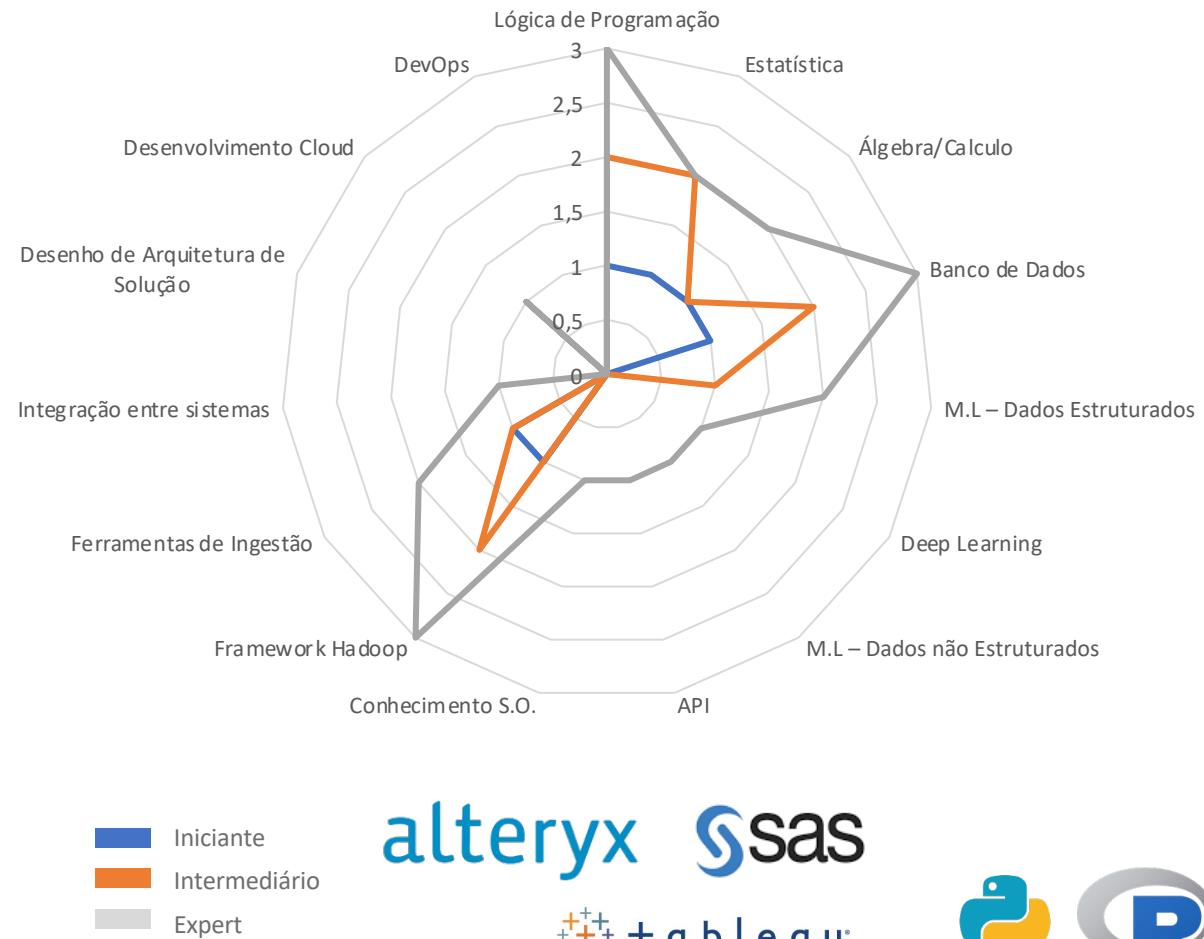


Profissionais x Skills	Data Science	Mathematics	Machine Learning	Algorithmic coding	Software eng.	Communication	Business acumen
Analista de Dados	✓			✓		✓	✓
Eng. Machine Learning	✓	✓	✓	✓	✓		
Eng. Software - ML			✓	✓	✓	✓	
Researcher ML		✓	✓	✓	✓		Research
Eng. de Software				✓	✓	✓	

O Perfil do Analista de Dados

O perfil é selecionado em entrevista técnica, análise de portfólio (Github), e seleção de perfil (gestor funcional).

PROCESSOS	Iniciante	Intermed.	Expert
Conhecimento			
Conhecimento Técnicos			
Lógica de Programação	1	2	3
Estatística	1	2	2
Álgebra/Calculo	1	1	2
Banco de Dados	1	2	3
Machine Learning – Dados Estruturados	-	1	2
Deep Learning	-	-	1
Machine Learning – Dados não Estruturados	-	-	1
Construção e Funcionamento de API	-	-	1
Conhecimento S.O.	-	-	1
Framework Hadoop	1	2	3
Ferramentas de Ingestão	1	1	2
Integração entre sistemas	-	-	1
Desenho de Arquitetura de Solução	-	-	-
Desenvolvimento Cloud	-	-	1
DevOps	-	-	-
Habilidades			
Solução de Problemas	1	2	3
Comunicação	1	2	3
Transformação/Inovação	1	2	2
Visão Estratégica e Sistêmica	1	2	3
Influência	-	-	1
Tutoria	-	2	3



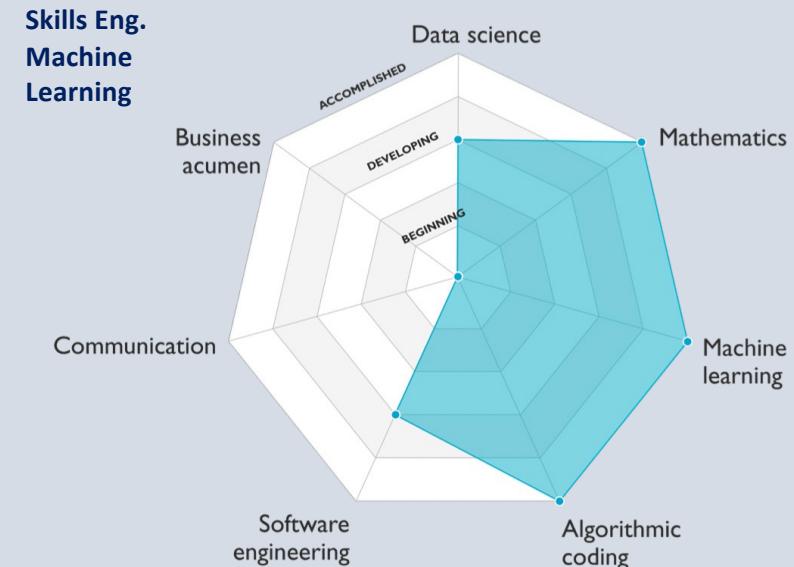
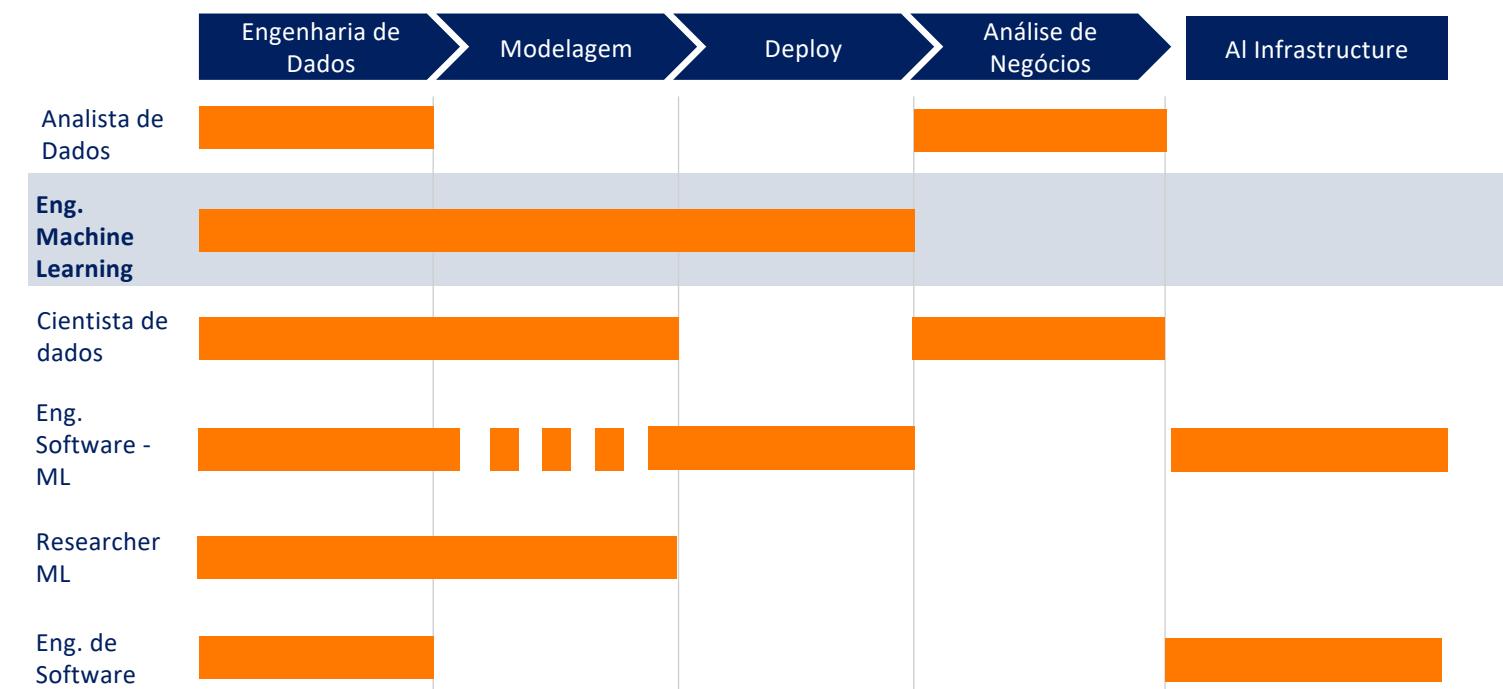
alteryx sas

tableau
SOFTWARE



Perfil do Eng. De Machine Learning

Frentes de trabalho x profissionais , de acordo com texto publicado no Workera:

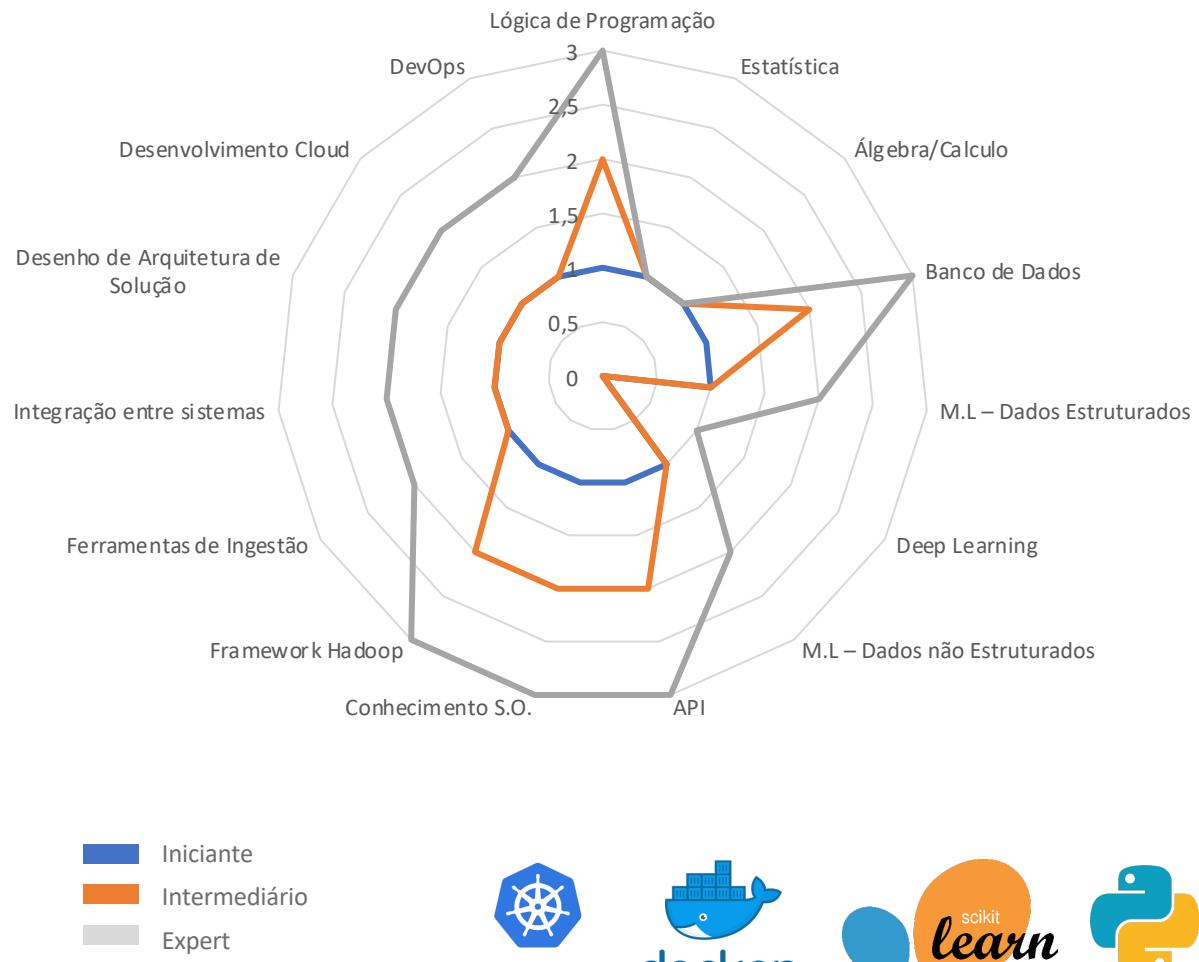


Profissionais x Skills	Data Science	Mathematics	Machine Learning	Algorithmic coding	Software eng.	Communication	Business acumen	Research
Analista de Dados	✓			✓		✓	✓	
Eng. Machine Learning	✓	✓	✓	✓	✓			
Eng. Software - ML			✓	✓	✓	✓		
Researcher ML		✓	✓	✓	✓			
Eng. de Software				✓		✓		

O Perfil do Engenheiro de Machine Learning

O perfil é selecionado em entrevista técnica, análise de portfólio (Github), e seleção de perfil (gestor funcional).

PROCESSOS			
Conhecimento	Iniciante	Intermed.	Expert
Conhecimento Técnicos			
Lógica de Programação	1	2	3
Estatística	1	1	1
Álgebra/Calculo	1	1	1
Banco de Dados	1	2	3
Machine Learning – Dados Estruturados	1	1	2
Deep Learning	-	-	1
Machine Learning – Dados não Estruturados	1	1	2
Construção e Funcionamento de API	1	2	3
Conhecimento S.O.	1	2	3
Framework Hadoop	1	2	3
Ferramentas de Ingestão	1	1	2
Integração entre sistemas	1	1	2
Desenho de Arquitetura de Solução	1	1	2
Desenvolvimento Cloud	1	1	2
DevOps	1	1	2
Habilidades			
Solução de Problemas	1	2	3
Comunicação	1	2	3
Transformação/Inovação	1	2	2
Visão Estratégica e Sistêmica	1	2	3
Influência	-	-	1
Tutoria	-	2	3

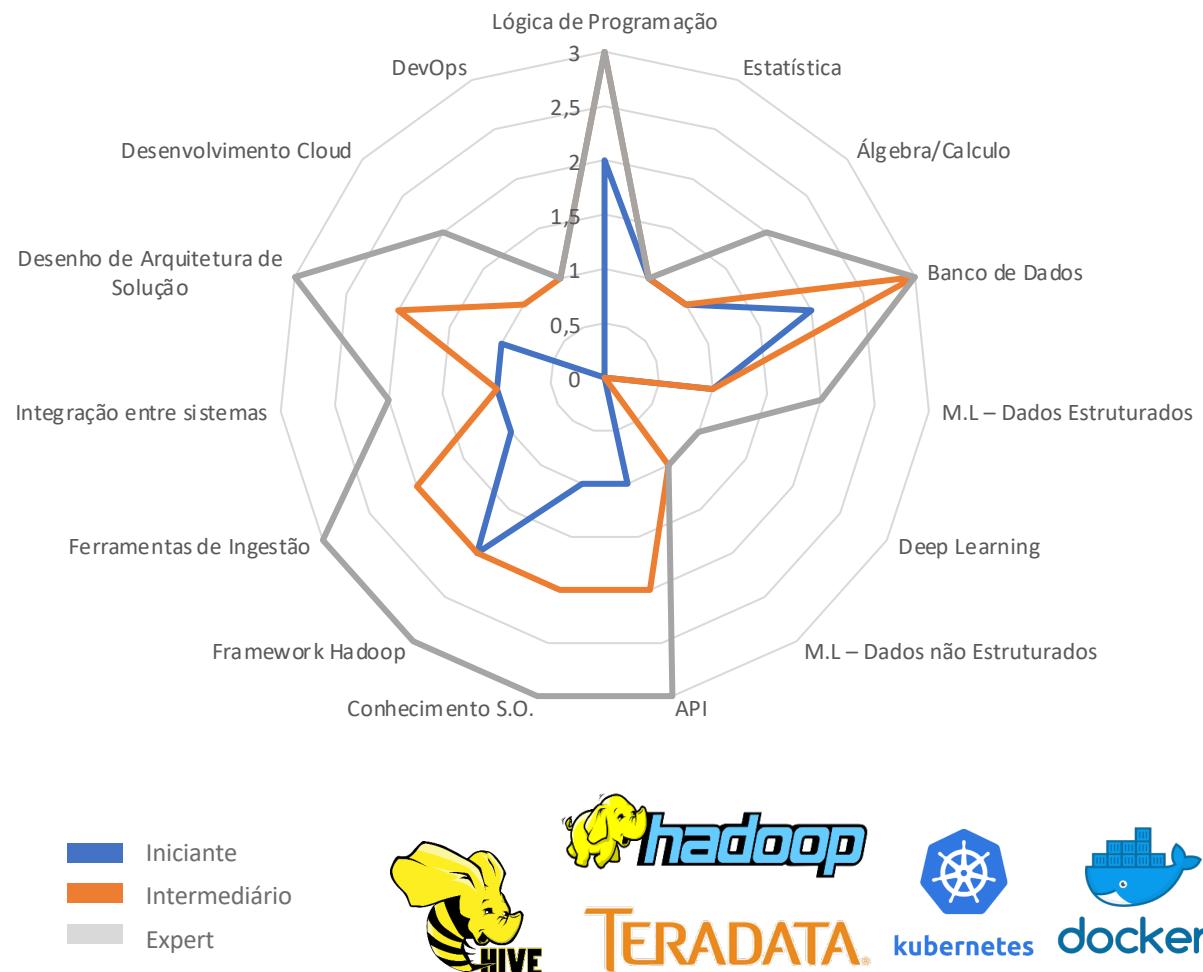


O Perfil do Engenheiro de Dados

O perfil é selecionado em entrevista técnica, análise de portfólio (Github), e seleção de perfil (gestor funcional).

PROCESSOS

	Iniciante	Intermed.	Expert
Conhecimento			
Conhecimento Técnicos			
Lógica de Programação	2	3	3
Estatística	1	1	1
Álgebra/Calculo	1	1	2
Banco de Dados	2	3	3
Machine Learning – Dados Estruturados	1	1	2
Deep Learning	-	-	1
Machine Learning – Dados não Estruturados	-	1	1
Construção e Funcionamento de API	1	2	3
Conhecimento S.O.	1	2	3
Framework Hadoop	2	2	3
Ferramentas de Ingestão	1	2	3
Integração entre sistemas	1	1	2
Desenho de Arquitetura de Solução	1	2	3
Desenvolvimento Cloud	-	1	2
DevOps	-	1	1
Habilidades			
Solução de Problemas	1	2	3
Comunicação	2	2	3
Transformação/Inovação	1	2	2
Visão Estratégica e Sistêmica	1	2	3
Influência	-	1	2
Tutoria	-	2	3



Como diabos eu faço para aprender isso?

- Pré-requisitos
 - Cálculo
 - Álgebra Linear
 - Estatística
 - Programação
- Livros
 - Artificial Intelligence: A Modern Approach
 - An Introduction to Statistical Learning
 - Hands-On Machine Learning with Scikit-Learn and TensorFlow
- Cursos
 - KhanAcademy.org
 - Stanford Statistical Learning (online)



- 
- 1** Motivação e **Contextualização Geral**
 - 2** Ciclo de vida do **dado**
 - 3** Ciclo de vida do **modelo**
 - 4** Técnicas e **Algoritmos**
 - 5** Perfil e responsabilidades dos **profissionais de dados**
 - 6** Aplicação e **Cases**

Making a business case for AI

- Who is the customer? Who is the user?
- **What is the business problem?**
- **Why use machine learning?**
- Does ML improve an existing use case or is this a new use case?
- **What are the data sources? Do you have right datasets?**
- What are the ethical, legal, or compliance issues?
- **Will this solution generate revenue? Save cost? Improve safety? Improve operations?**
- How will you maintain the solution?



Use case



Use case / business

- Business problem
- User, customer, actor
 - ✓ Flow of events
 - ✓ Flow of alternate events
- Functionality and result
- Benefit
- Process change

Use case / ML technical

- Describe how ML improves the existing use case
- Focus on what to predict and signals
- Describe signals, data sources, dataset
- Specify how to measure results

Success criteria

What the customer wants



Reduce risk

A better understanding of what is happening to get more visibility

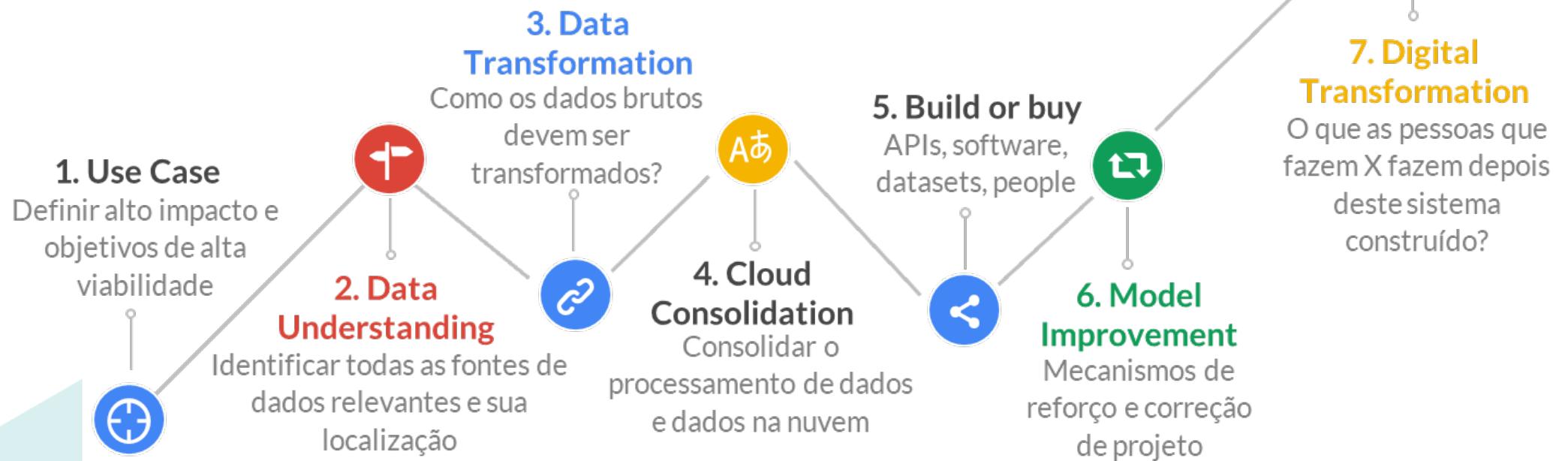
Optimize costs

More output for lower cost through understanding of the value chain and waste drivers

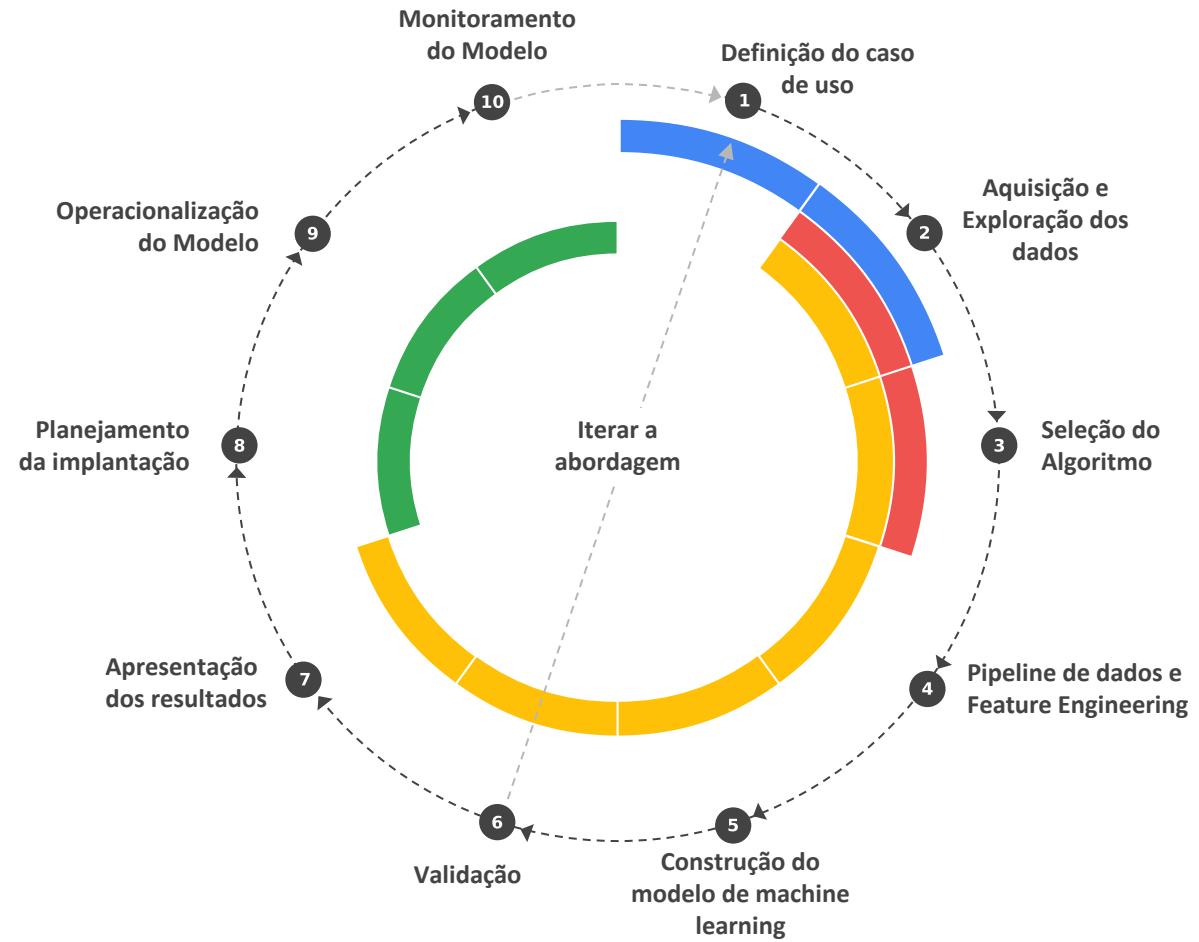
Grow

A better user or customer experience to increase usage and adoption of a product

ML Journey



ML lifecycle



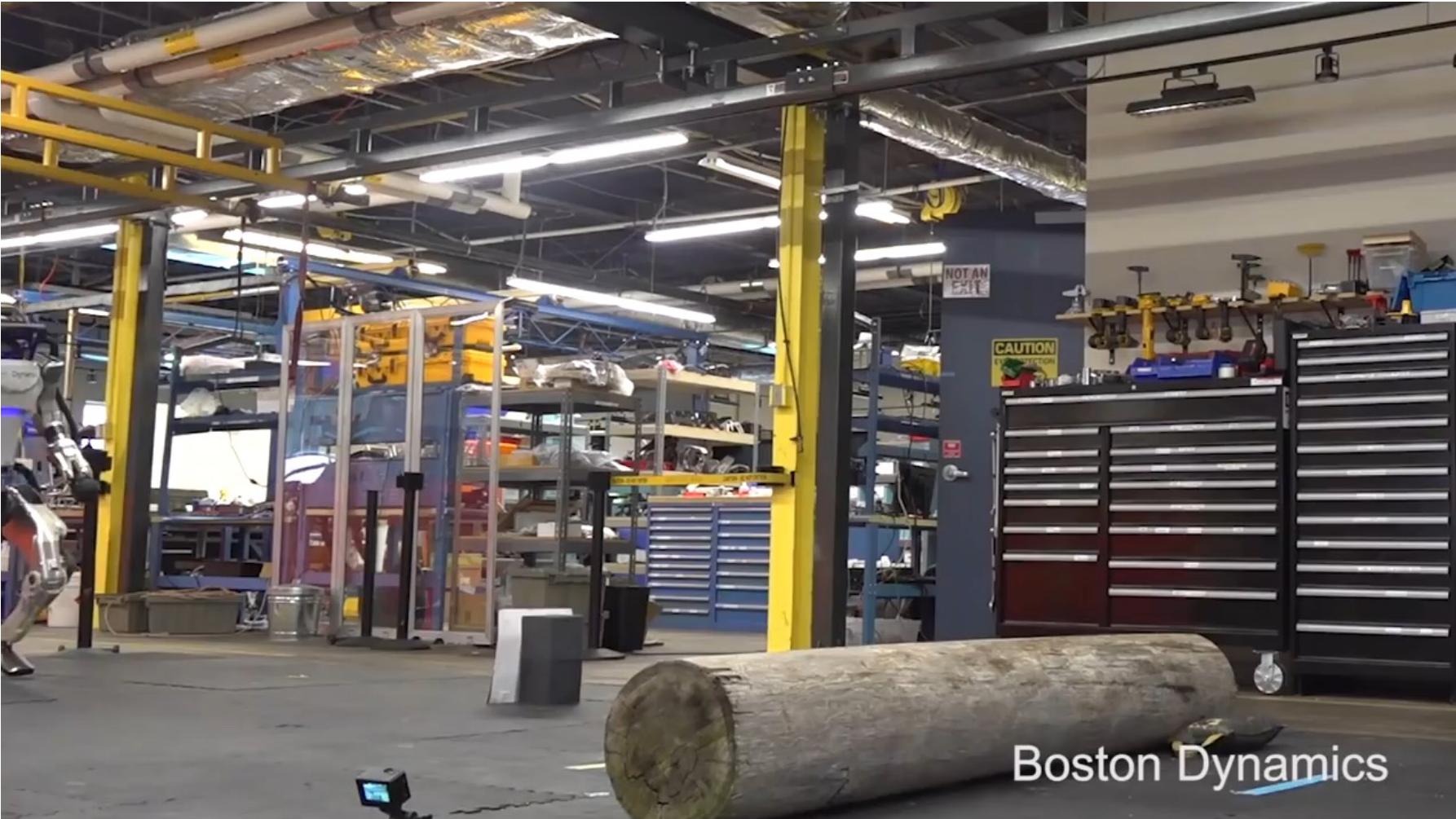
Realidade: ML requer DevOps



ROBÔ E OBAMA



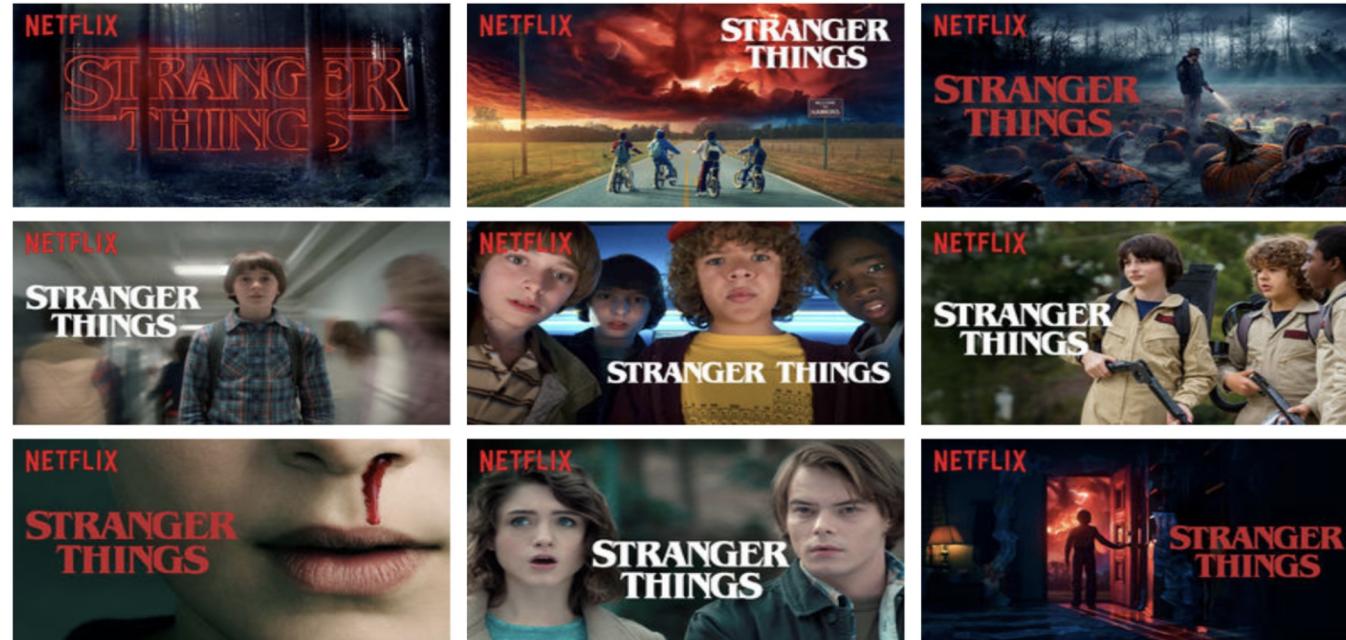
ATLAS, DA BOSTON DYNAMICS



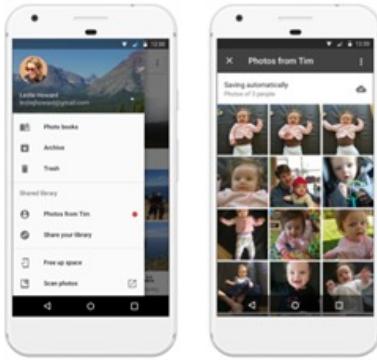
Boston Dynamics



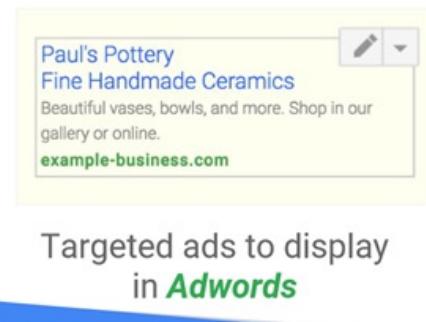
Media & Entertainment: Artwork Personalization



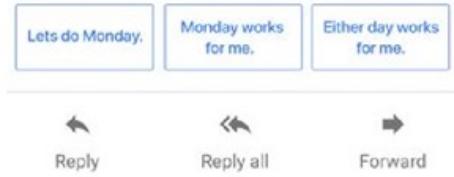
Artwork for Stranger Things that each receive over 5% of impressions from our personalization algorithm.
Different images cover a breadth of themes in the show to go beyond what any single image portrays.



Classify pictures in
Google Photos



Targeted ads to display
in **Adwords**



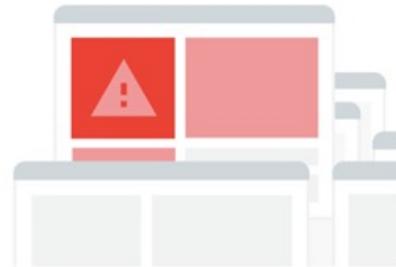
Smart reply in **Inbox**



Recommendations for the
next video in **Youtube**



Pedestrian detection
Self-driving cars



Spam detection in **Gmail**



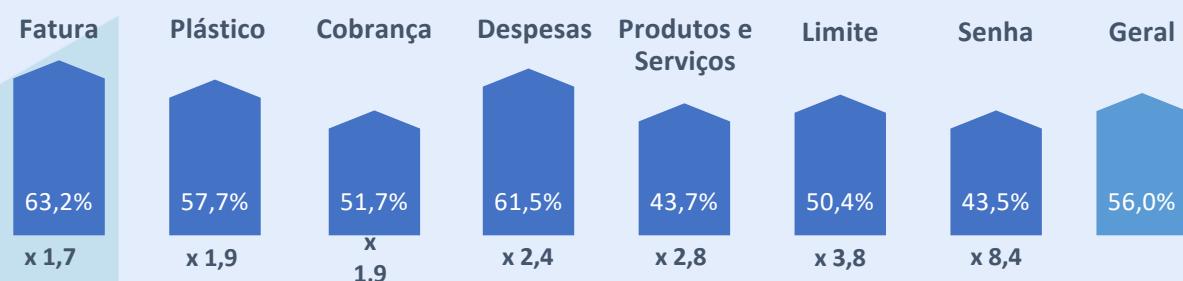
Reconhecimento inteligente de demanda



Conseguimos **prever o motivo da ligação** com base na **jornada do cliente**?



Resultados (por macroassunto)



Objetivo

Atender o cliente de **forma proativa e mais assertiva**

Ações com base na saída do modelo

Antes do contato

- Envio de SMS
- Push App
- Ligação Ativa
- Backoffice

Durante o contato

- Contextualização

Impacto (antes do contato)

Potencial de evitar
720 mil ligações por mês

Futuro

- PF
- AVI
- PJ



Análise facial



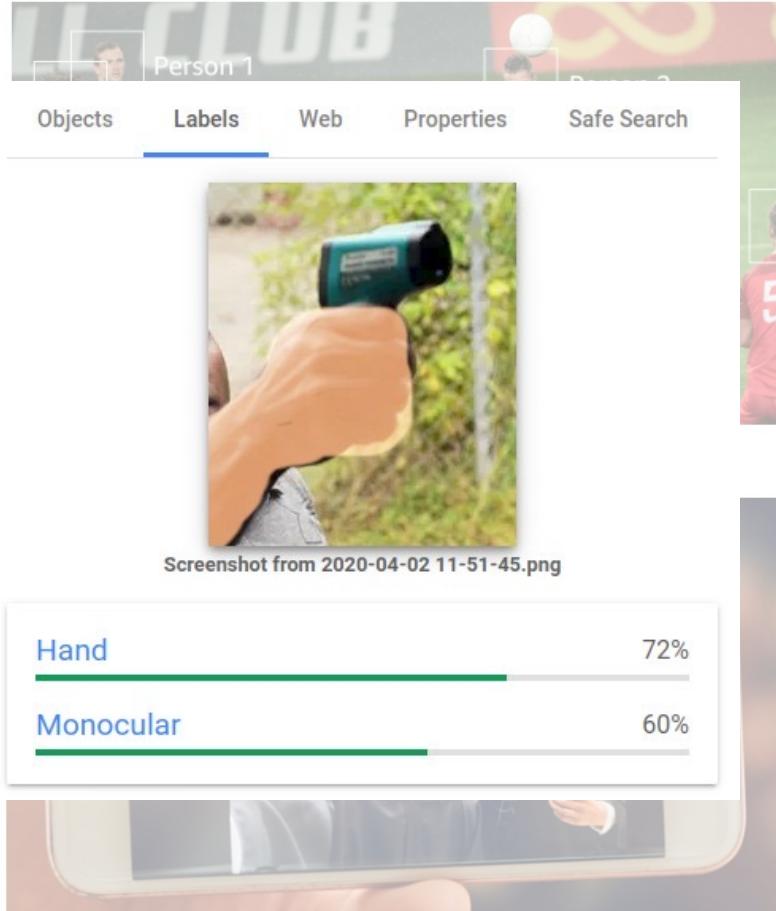
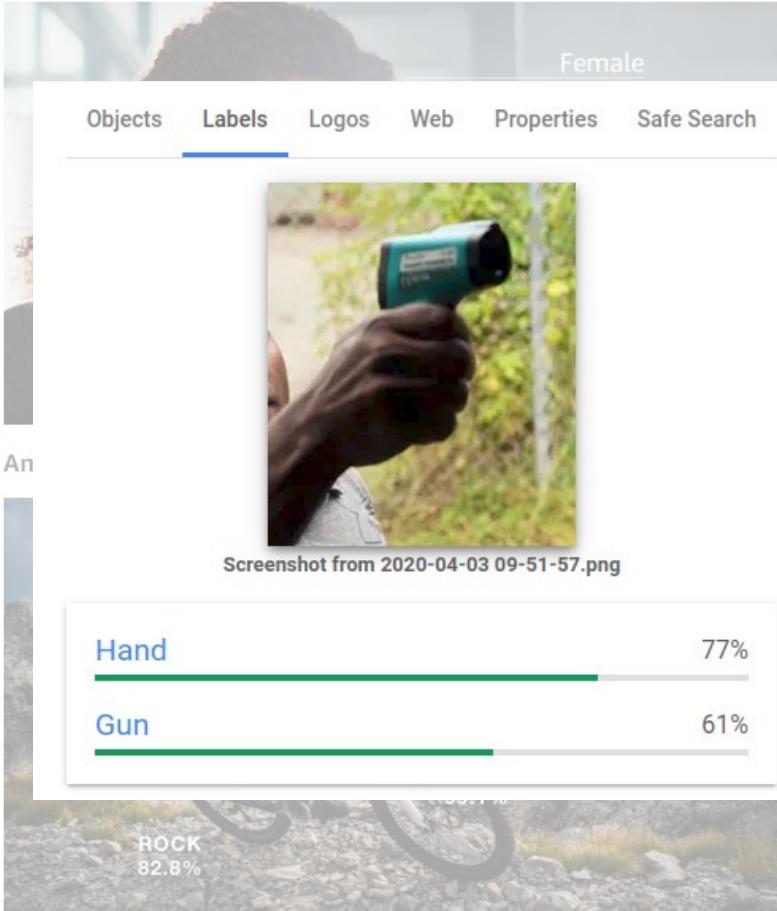
Determinação de caminhos



Detecção de objetos, cenas e atividades



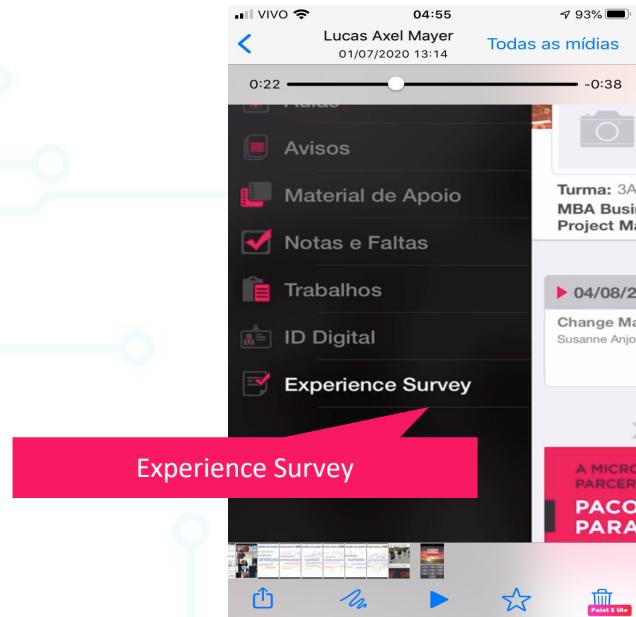
Reconhecimento facial



O que você achou da aula de hoje?

Pelo aplicativo da FIAP

(Entrar no FIAPP, e no menu clicar em Experience Survey)



Pelo QR Code



Pelo link

<https://fiap.me/Pesquisa-MBA>

Obrigado!

profdheny.fernandes@fiap.com.br



FIAP MBA⁺

Copyright © 2018 | Professor Dheny R. Fernandes

Todos os direitos reservados. Reprodução ou divulgação total ou parcial deste documento, é expressamente proibido sem consentimento formal, por escrito, do professor/autor.