

# Introdução ao **Aprendizado de Máquina** com **scikit-learn**

MAGRATHEA LABS

[www.magrathealabs.com](http://www.magrathealabs.com)

WHEN A USER TAKES A PHOTO,  
THE APP SHOULD CHECK WHETHER  
THEY'RE IN A NATIONAL PARK...

SURE, EASY GIS LOOKUP.  
GIMME A FEW HOURS.

... AND CHECK WHETHER  
THE PHOTO IS OF A BIRD.

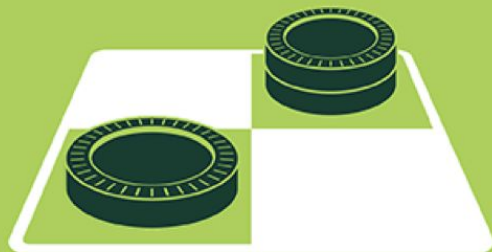
I'LL NEED A RESEARCH  
TEAM AND FIVE YEARS.



IN CS, IT CAN BE HARD TO EXPLAIN  
THE DIFFERENCE BETWEEN THE EASY  
AND THE VIRTUALLY IMPOSSIBLE.

# ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



## MACHINE LEARNING

Machine learning begins to flourish.



## DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

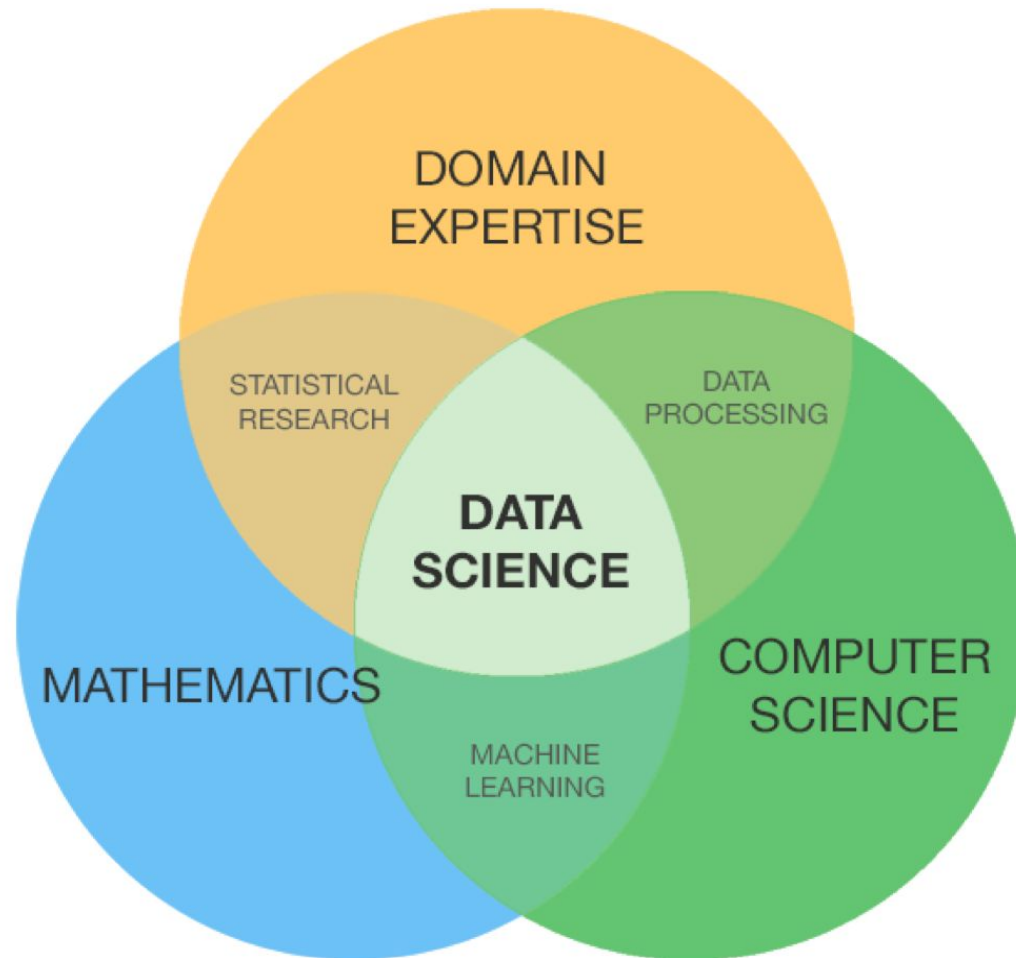
1980's

1990's

2000's

2010's

Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



*Source: Palmer, Shelly. Data Science for the C-Suite.  
New York: Digital Living Press, 2015. Print.*





# O INVERNO DA INTELIGÊNCIA ARTIFICIAL

# Machine Learning

"Machine learning é a extração semi-automatizada de conhecimento a partir de dados"

# Machine Learning (aprendizado de máquina)

Popularidade recente devido à disponibilidade de:

- Dados;
- Poder computacional;
- Boas ferramentas/bibliotecas/técnicas.

# Machine Learning

Aprendizado  
Supervisionado

Orientado à tarefa  
Classificação  
Regressão

Aprendizado  
Não-supervisionado

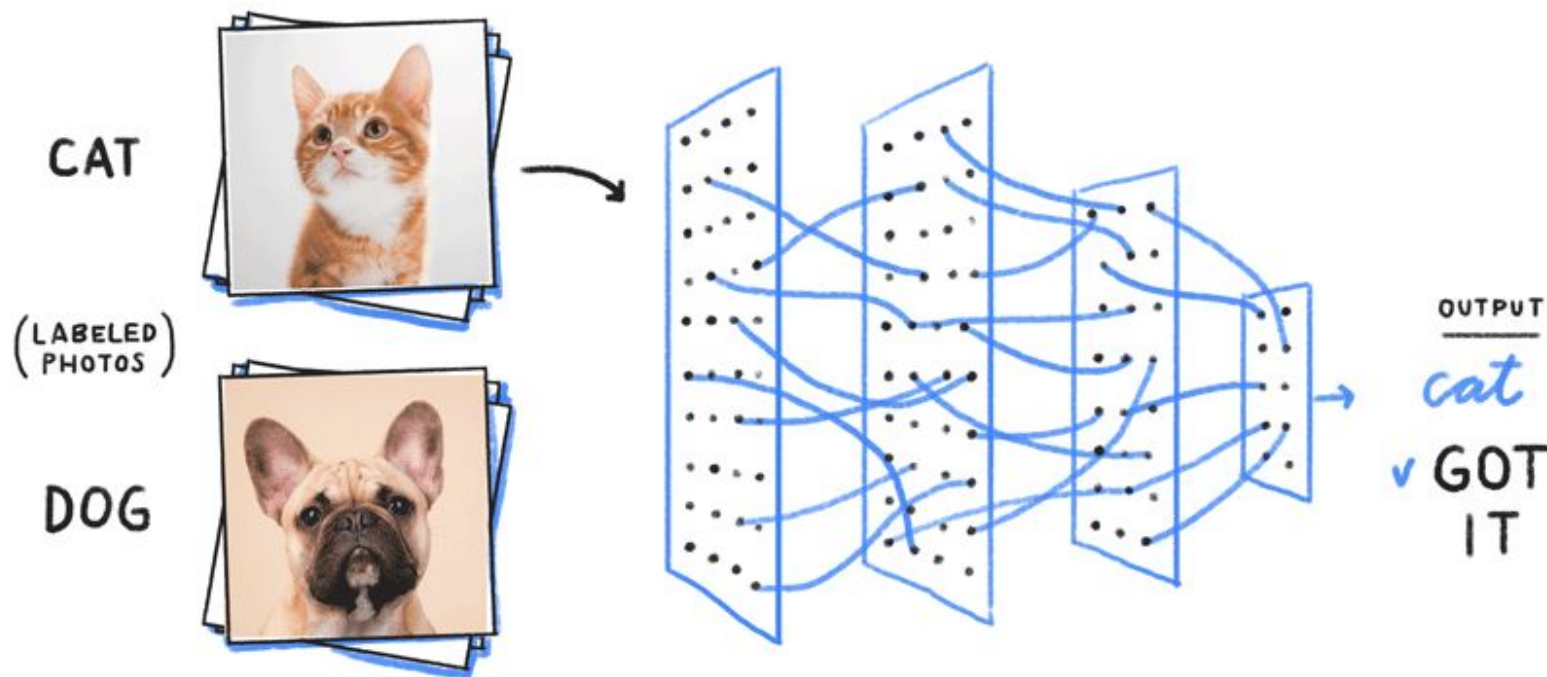
Orientado aos dados  
Clustering

Aprendizado  
por Reforço

Orientado ao ambiente  
Agentes Inteligentes



# Aprendizado Supervisionado

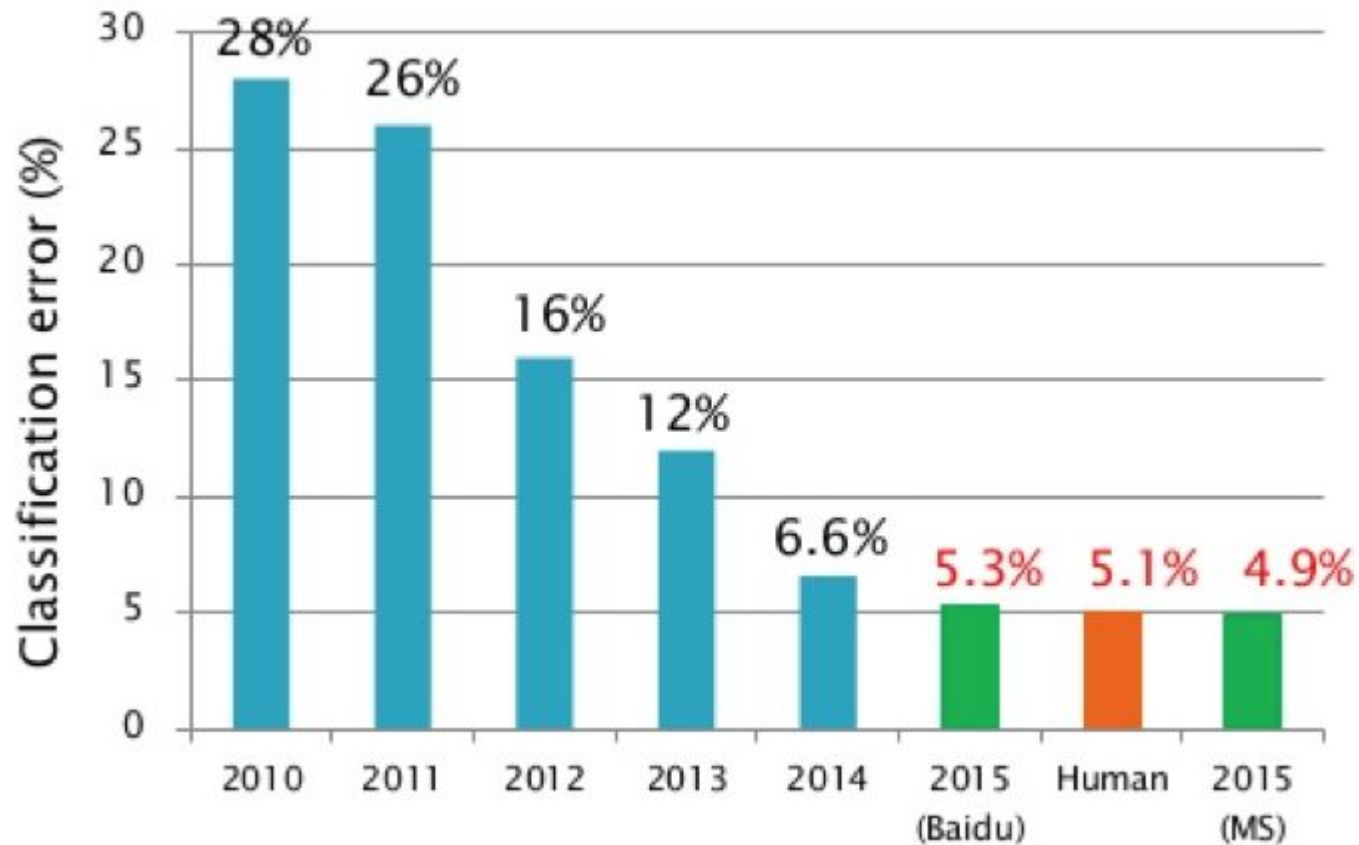


# Aprendizado Não-supervisionado



# Aprendizado por Reforço





- SIFT (Scale Invariant Feature Transformation) + SVM
- Problema: Como extrair características das imagens?
- 2012: GPUs
- Grandes Bases de dados ao invés de extração de características





UM EXEMPLO REAL

## Heineken - Brand

Search by Keyword or Brand Name

Heineken

Sort by Relevance

Time Frame

Brand

Object

Scene Type

Location

Content

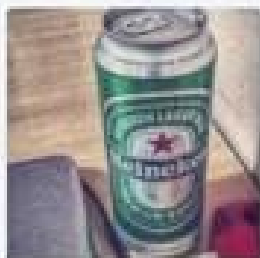
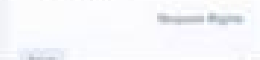
Size

Reset All Filters

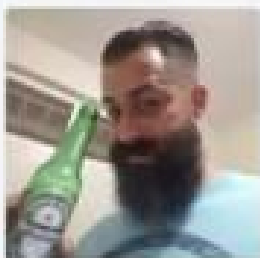
Download



**Foto van de Agt Landen**  
© Heineken Nederland - Sep 1985, 2015  
Heineken Nederland  
100% Heineken



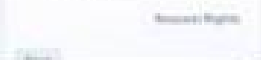
**Aling mark**  
© Heineken Nederland - Sep 1985, 2015  
Heineken Nederland  
100% Heineken



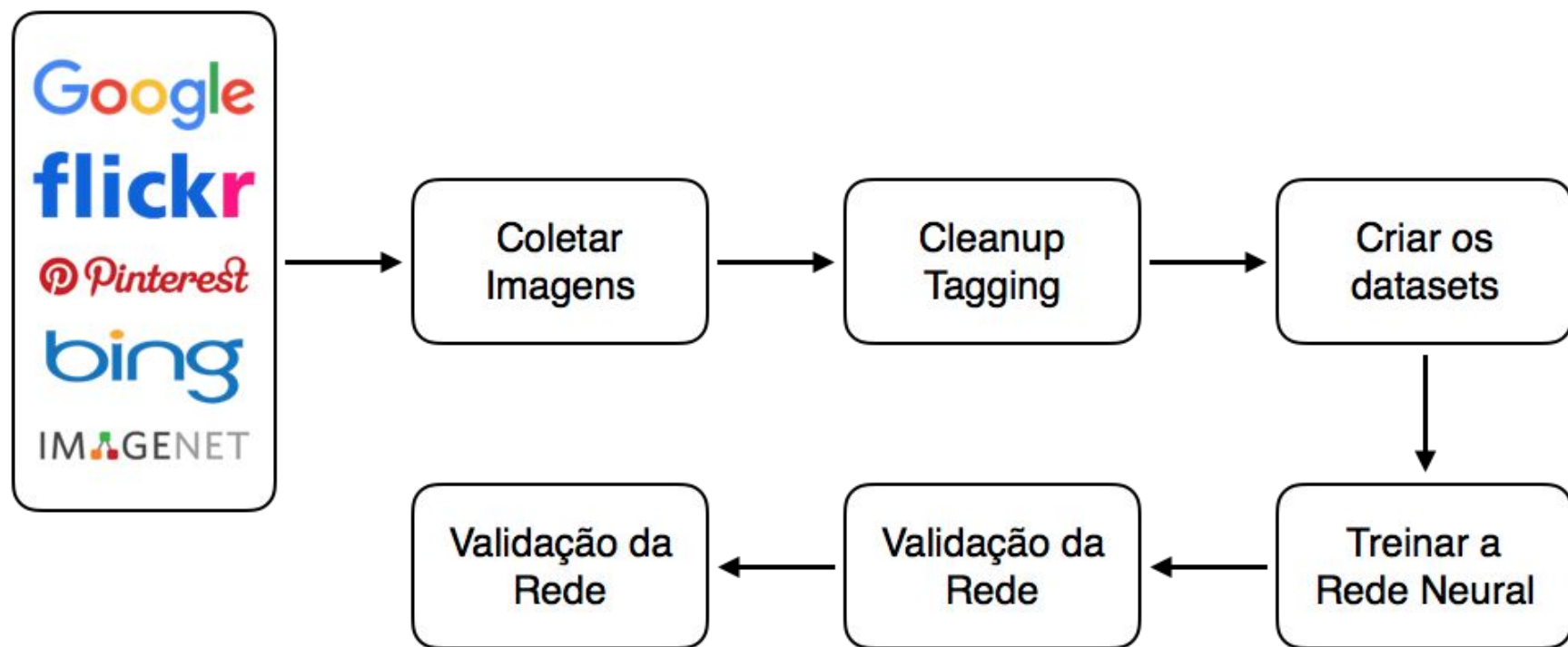
**We Are Heineken Beards**  
© Heineken Nederland - Sep 1985, 2015  
Heineken Nederland  
100% Heineken



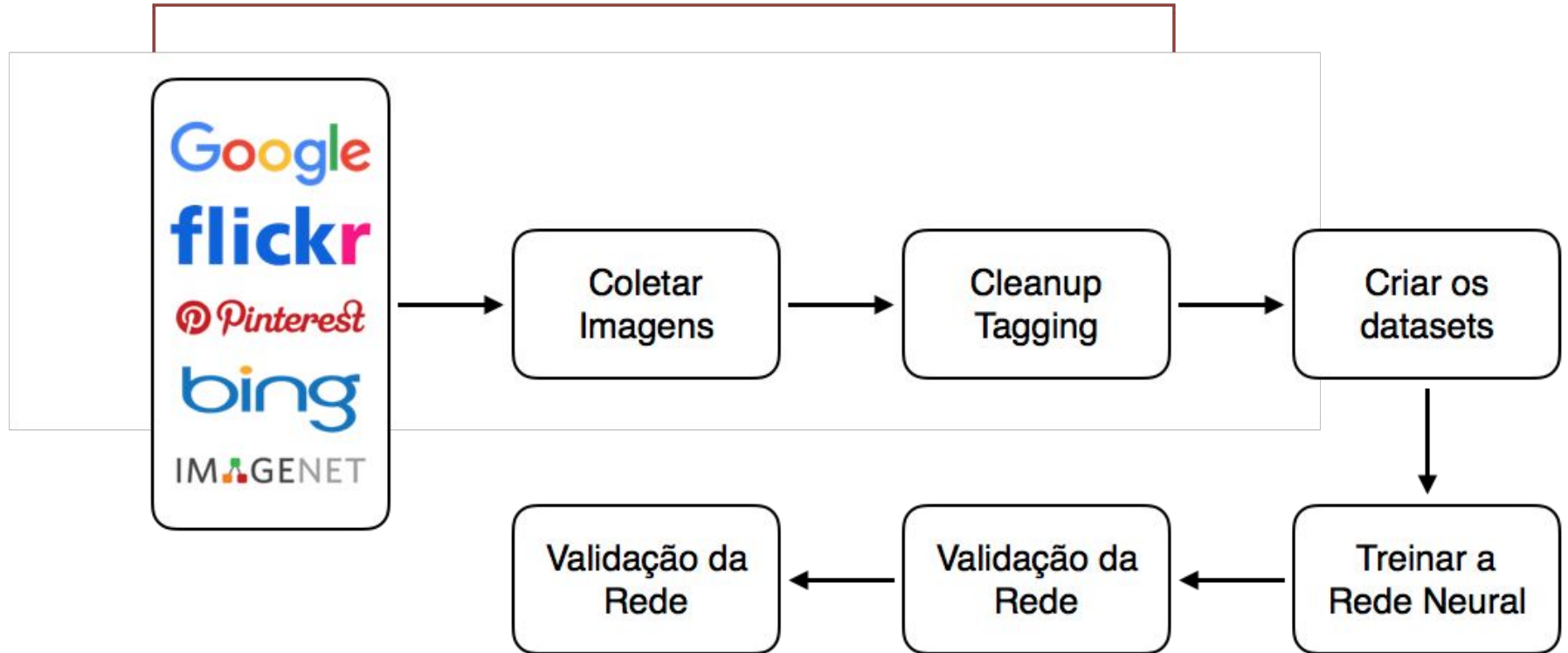
**Beerdjes River**  
© Heineken Nederland - Sep 1985, 2015  
Heineken Nederland  
100% Heineken

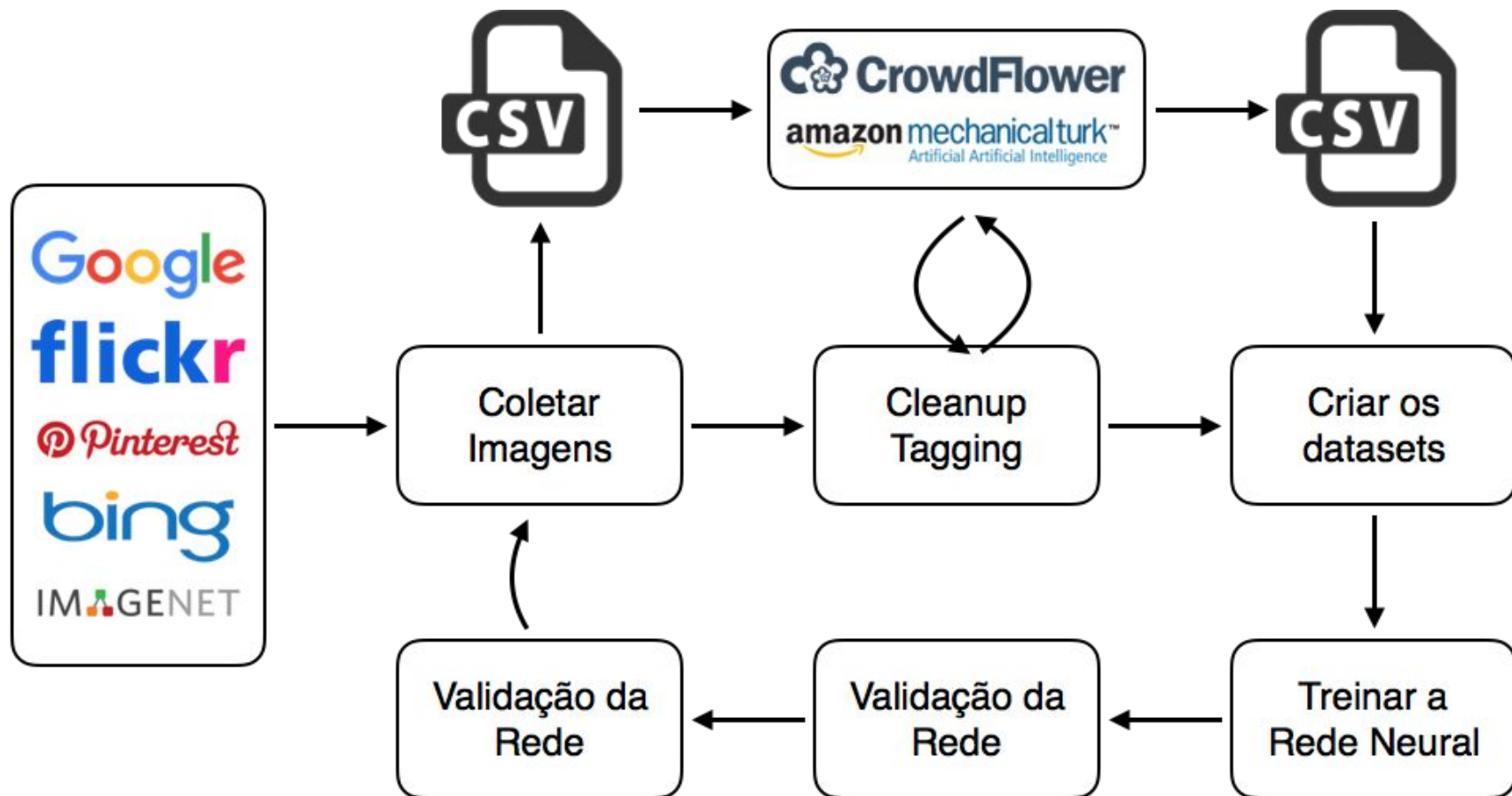




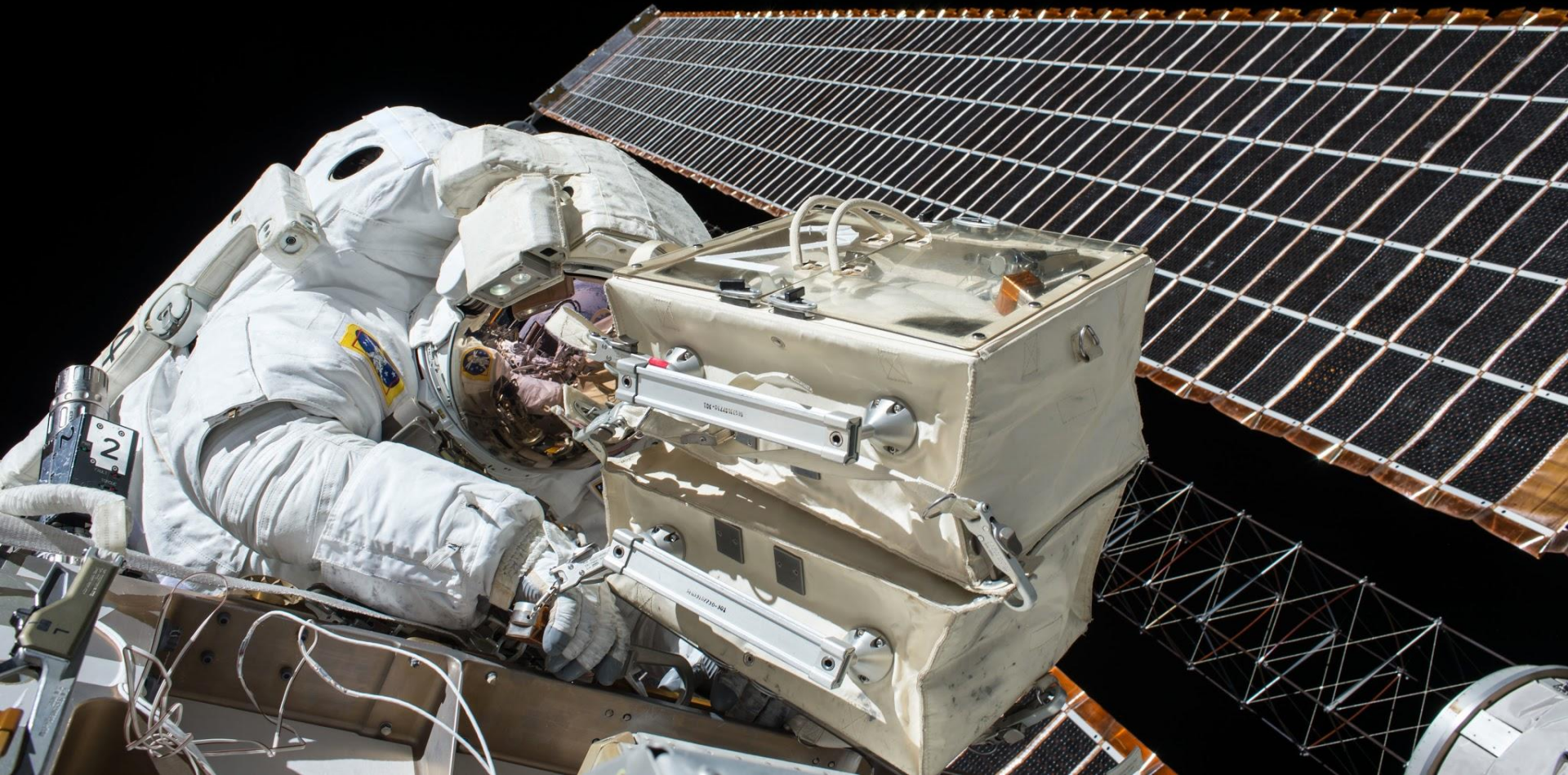


90/95% do tempo





# DESAFIOS



# DESAFIOS

Necessidade de:

- Profissionais qualificados
- Poder computacional
- Disponibilidade de dados “prontos para uso”
- Reconhecer a aplicabilidade da tecnologia



# COMO APLICAR MACHINE LEARNING?

1. Determine objetivos, métricas e restrições;
2. Avalie os dados disponíveis;
3. Treine um modelo (algoritmo);
4. Integre os modelos e realize testes;
5. Monitore a operação do modelo;
6. Volte ao passo 1.

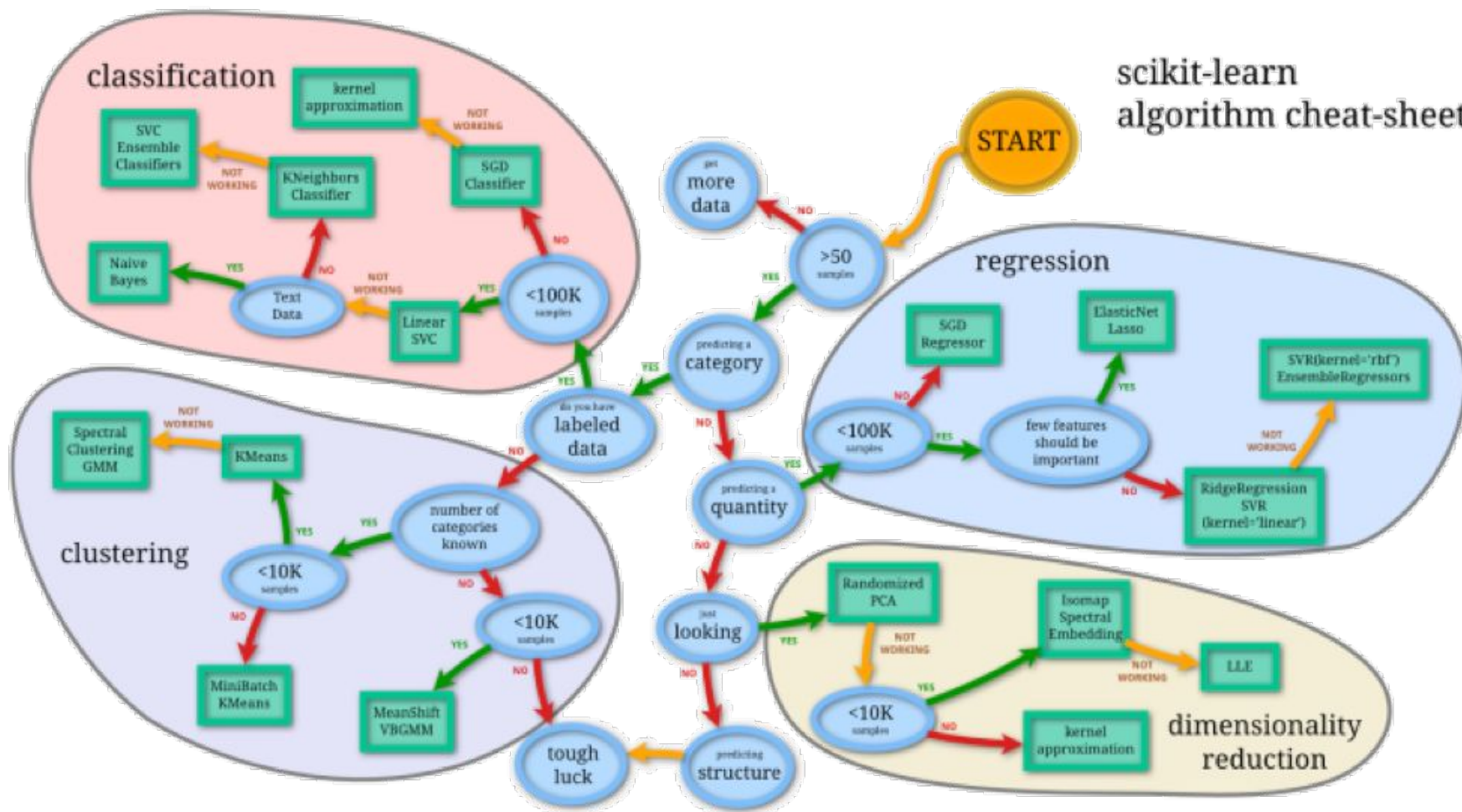


# QUESTÕES

- Como escolher os atributos do conjunto de dados?
- Qual modelo eu uso?
- Como otimizar o modelo para melhorar o desempenho?
- Como garantir que o modelo vai generalizar para dados não vistos?
- Posso estimar quão bem meu modelo vai desempenhar para novos dados?

# scikit-learn

- Interface consistente para os modelos disponíveis;
- Vários parâmetros de configuração;
- Bons valores padrões para os parâmetros;
- Ótima documentação;
- Batteries included;
- Comunidade ativa de devs e suporte.

scikit-learn  
algorithm cheat-sheet

DEMO

# Joinville Machine Learning

Comunidade de empreendedores, engenheiros, cientistas e entusiastas de Aprendizado de Máquinas de Joinville, SC



## Comunidade

Participe do **Clube de Computação Aplicada** que acontece semanalmente na quarta-feira a partir das 19 horas no [Fab Lab Joinville](#).

Este é o site da comunidade Joinville Machine Learning. Nossa comunidade nasceu depois que algumas pessoas se encontraram na Semana da Computação da UDESC. Realizamos um evento mensal para apresentar os projetos em que estamos trabalhando, artigos interessantes e tutoriais de ferramentas. Somos entusiastas, cientistas e engenheiros que estão atuando na área, estudando ou tentando descobrir um pouco mais sobre o assunto. *Não existe nenhum pré-requisito para participar dos nossos eventos.*

Participe do nosso [Meetup](#) e acompanhe os projetos da comunidade no [GitHub](#) e no [Kaggle](#).

## Últimos Eventos

Os materiais apresentados nos encontros estão no nosso [repositório de eventos](#).

- [17 de Maio \(Apresentações\)](#)
- [18 de Abril \(Apresentações\)](#)

# Clube de Computação Aplicada

## quarta-feira, 18 horas



[www.fablabjoinville.com.br](http://www.fablabjoinville.com.br)  
[github.com/fablabjoinville](https://github.com/fablabjoinville)



Conhece uma dessas paradas? Me manda seu CV:  
contato@magrathealabs.com  
[careers.jobscore.com/careers/magrathealabs](https://careers.jobscore.com/careers/magrathealabs)



MAGRATHEA LABS



MAGRATHEA LABS

+55 47 99155-2269

[www.magrathealabs.com](http://www.magrathealabs.com)

[contato@magrathealabs.com](mailto:contato@magrathealabs.com)