

Introdução à Inteligência Computacional

Apresentação da Disciplina

Cristiano Leite de Castro - crislcastro@ufmg.br

Objetivo do Curso

- Introduzir os **conceitos básicos** necessários para estudos na área de Inteligência Computacional.
 - caracterização e modelagem de problemas a partir de dados.
 - técnicas para análise, extração e seleção de atributos relevantes em conjuntos de dados;
 - modelos lineares e não-lineares para regressão, classificação e agrupamento.
 - técnicas para avaliação e seleção de modelos.

Aprendizado Estatístico

- Para introduzir esses conceitos, iremos utilizar um conjunto de ferramentas para **modelagem e análise de dados** denominado *Aprendizado Estatístico*;
- Abordagem estatística para o problema de *Aprendizado de Máquina*.

Breve Histórico

Neural Network Methods:

Statistical Learning Methods:

...

1940-50s McCulloch-Pits' neural model
Hebbian Learning Rule

1950-60s Rosenblatt's Perceptron

1960-70s Minsk and Papert's Book

1970-80s

1980-90s Hopfield and Kohonen NNs
BackPropagation (MLPs)

1990-00s Radial Basis Function NNs
CNNs

2000 - ...

Linha do tempo

Linear Regression
Fisher's Linear Discriminant Analysis
Logistic Regression

Hierarchical Clustering

K-means Algorithm
Vapnik and Chervonenkis' Theory

Generalized Linear Models

Classification and Regression Trees
Cross Validation
Ensemble Methods (Boosting)

Support Vector Machines

Convergência (Machine Learning)

Deep Learning, ELMs, Kernel Methods, LSTMs, Transformers,
etc.

Machine Learning

● machine learning
Search term

● big data
Search term

● data science
Search term

+ Add comparison

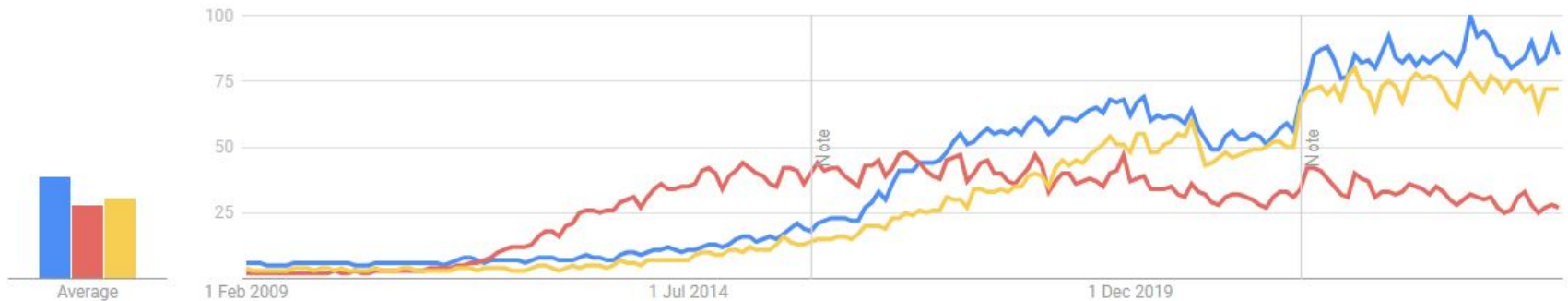
Worldwide ▼

18/02/2009 - 18/03/2025 ▼

All categories ▼

Web Search ▼

Interest over time (?)



O que é Machine Learning?

- Machine Learning (ML) is the study of computer algorithms that improve automatically through experience. [Mitchell, 1997]
- ML is an umbrella term for solving problems for which development of algorithms by human programmers would be cost-prohibitive, and instead the problems are solved by helping machines 'discover' their 'own' algorithms, without needing to be explicitly told what to do by any human-developed algorithms. [wikipedia]

Ethem Alpaydin (2020). Introduction to Machine Learning. Fourth ed. MIT.

Machine Learning

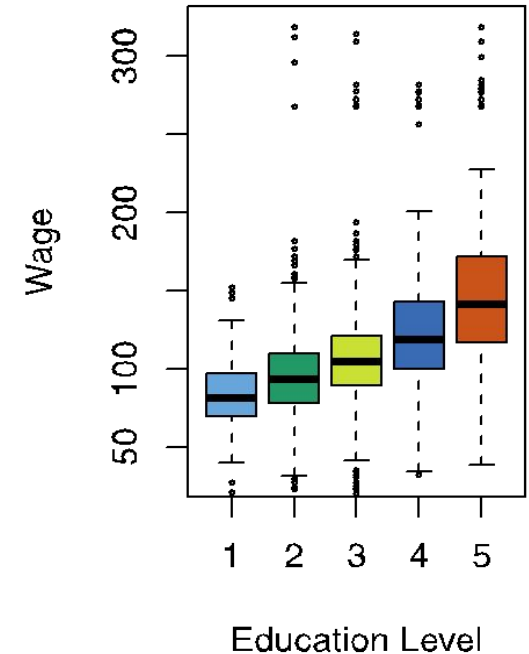
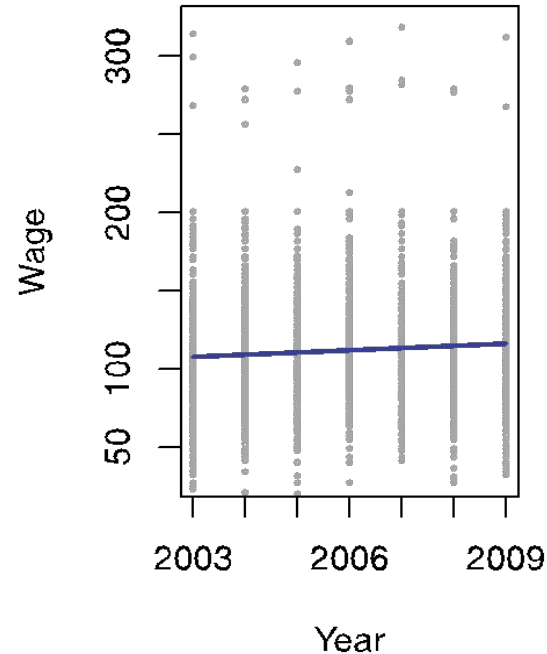
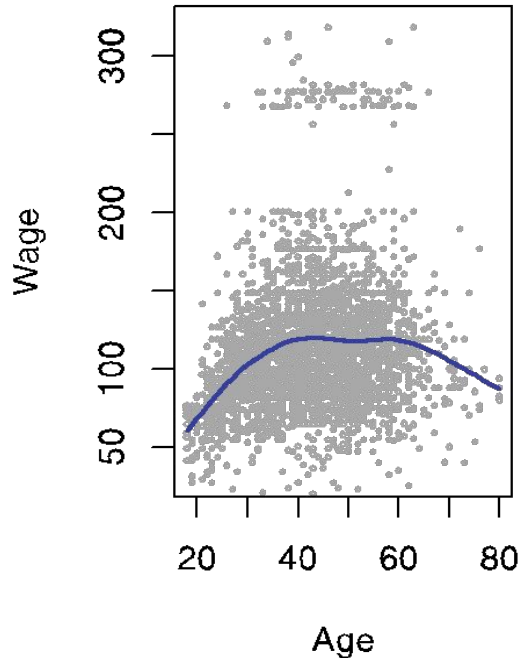
- **Premissa:**
 - Não existe um modelo (ferramenta) único que funcione bem para todos os problemas.
- Por isso, é importante entender as propriedades de cada modelo antes de aplicá-los aos dados, bem como saber sobre suas forças e fraquezas.

Machine Learning

- **Aprendizado Supervisionado:**
 - Dadas medições $(\mathbf{X}_1, Y_1), \dots, (\mathbf{X}_n, Y_n)$, aprender um modelo para **prever** Y_i baseado em \mathbf{X}_i
- **Aprendizado Não-Supervisionado:**
 - Dadas medições $(\mathbf{X}_1, \dots, \mathbf{X}_n)$, descobrir alguma **estrutura** com base em similaridade.

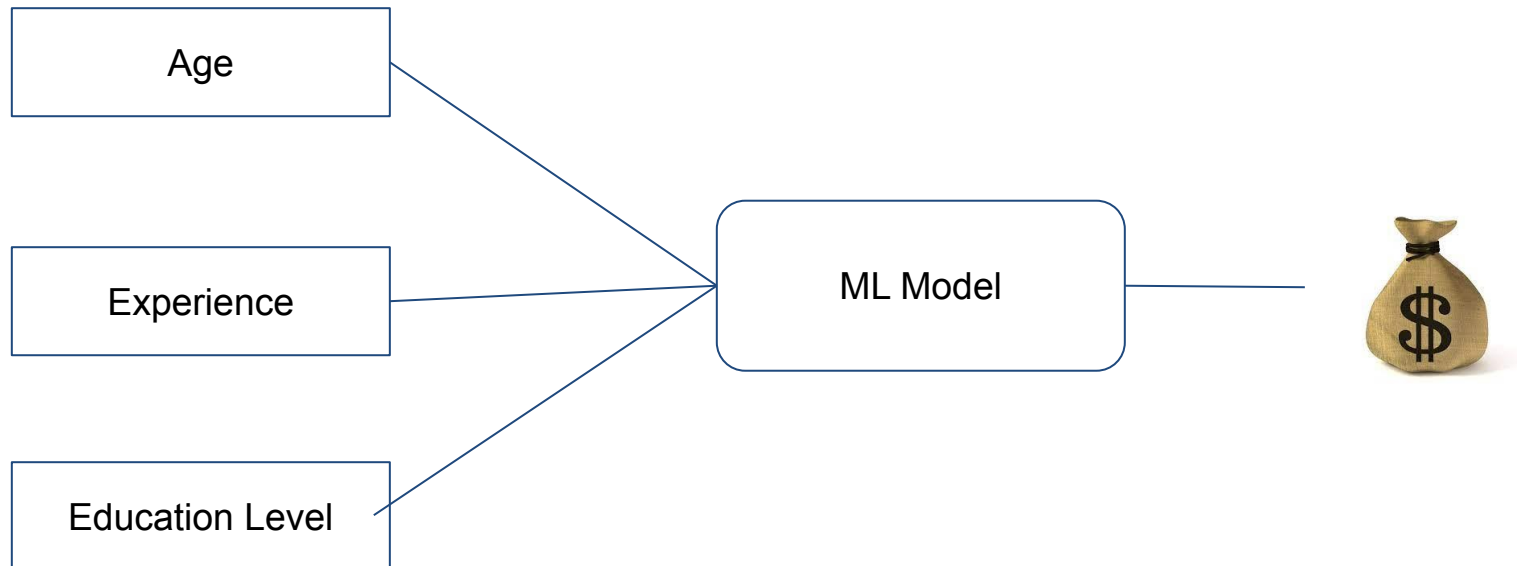
Aprendizado Supervisionado: Exemplo

- *Wage data*: compreender a relação entre idade, experiência e nível educacional, com o salário recebido.



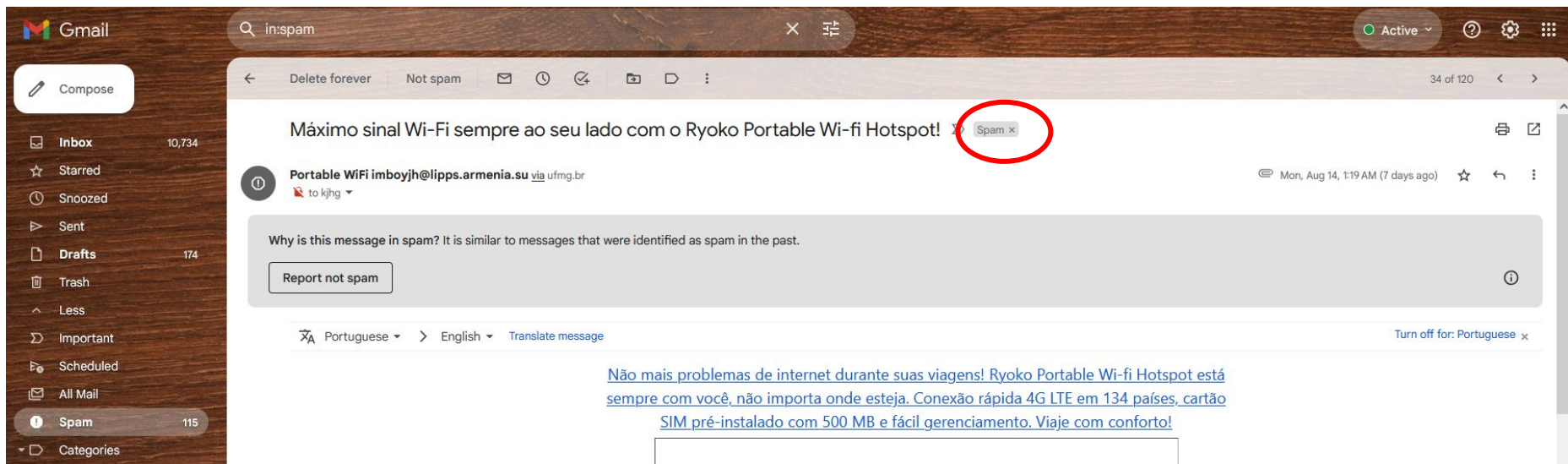
Aprendizado Supervisionado: Exemplo

- *Wage data*: compreender a relação entre idade, experiência e nível educacional, com o salário recebido.



Aprendizado Supervisionado: Exemplo

- **Filtro de spam:**



$\mathbf{X}_i \rightarrow$ email

$\mathbf{Y}_i \rightarrow$ spam ou não-spam (ham)

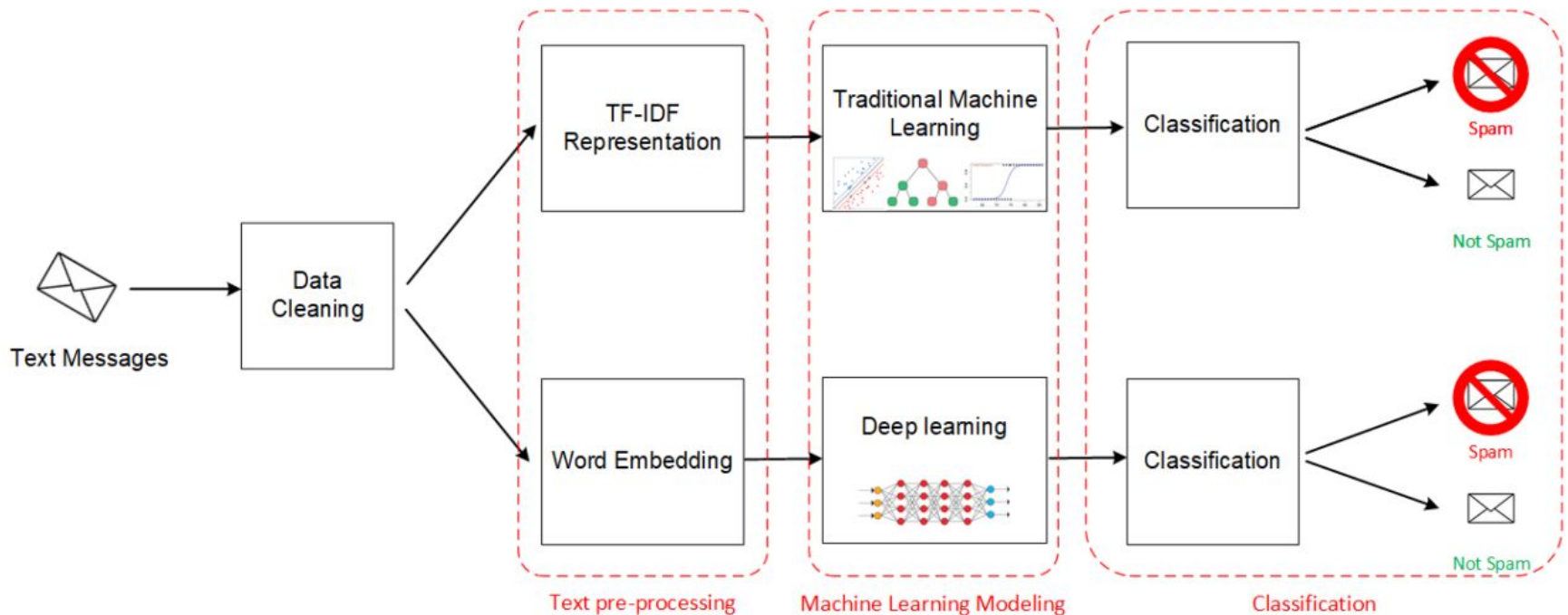
objetivo: prever \mathbf{Y}_i com base em \mathbf{X}_i

Aprendizado Supervisionado: Exemplo

- **Filtro de spam:**

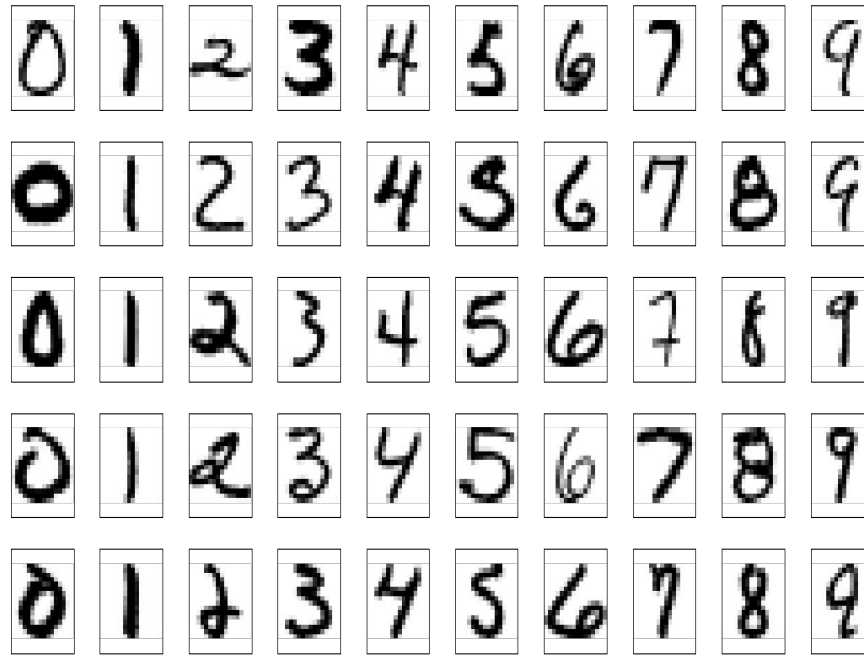
**Pré-processamento e Extração de Características
(embedding)**

Modelagem, Validação e Seleção de Modelos



Aprendizado Supervisionado: Exemplo

- Reconhecimento de dígitos escritos a mão

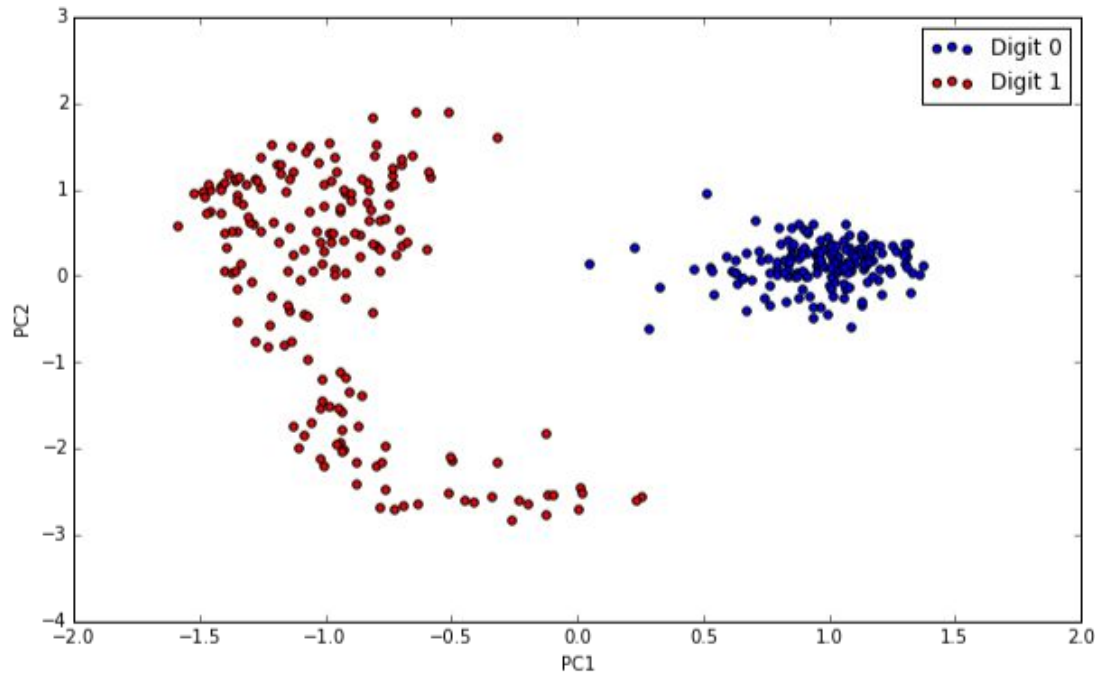
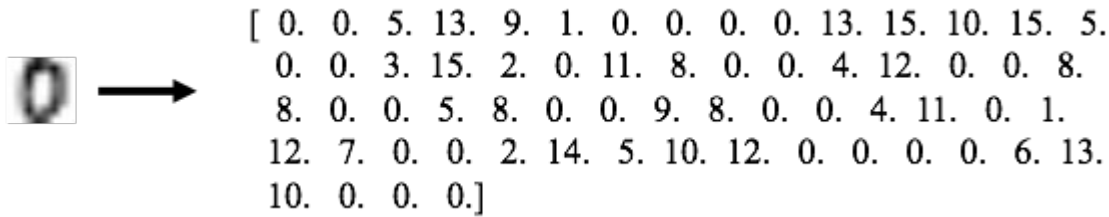


$X_i \rightarrow$ imagem (8x8 pixels)

$Y_i \rightarrow \{0, 1, \dots, 9\}$

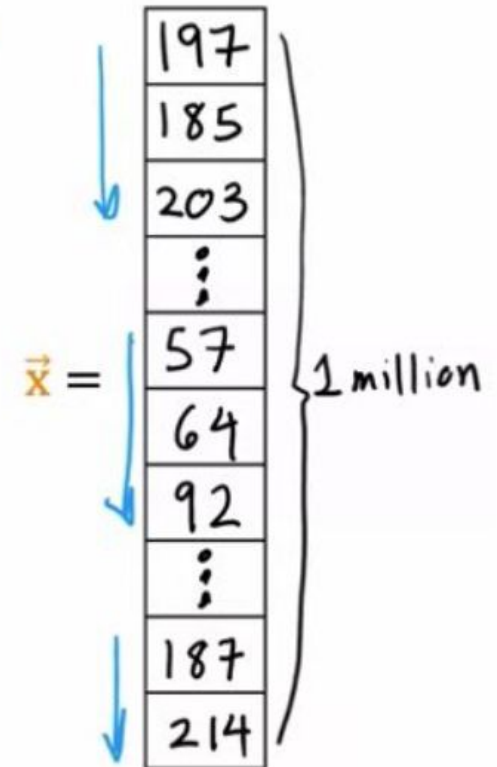
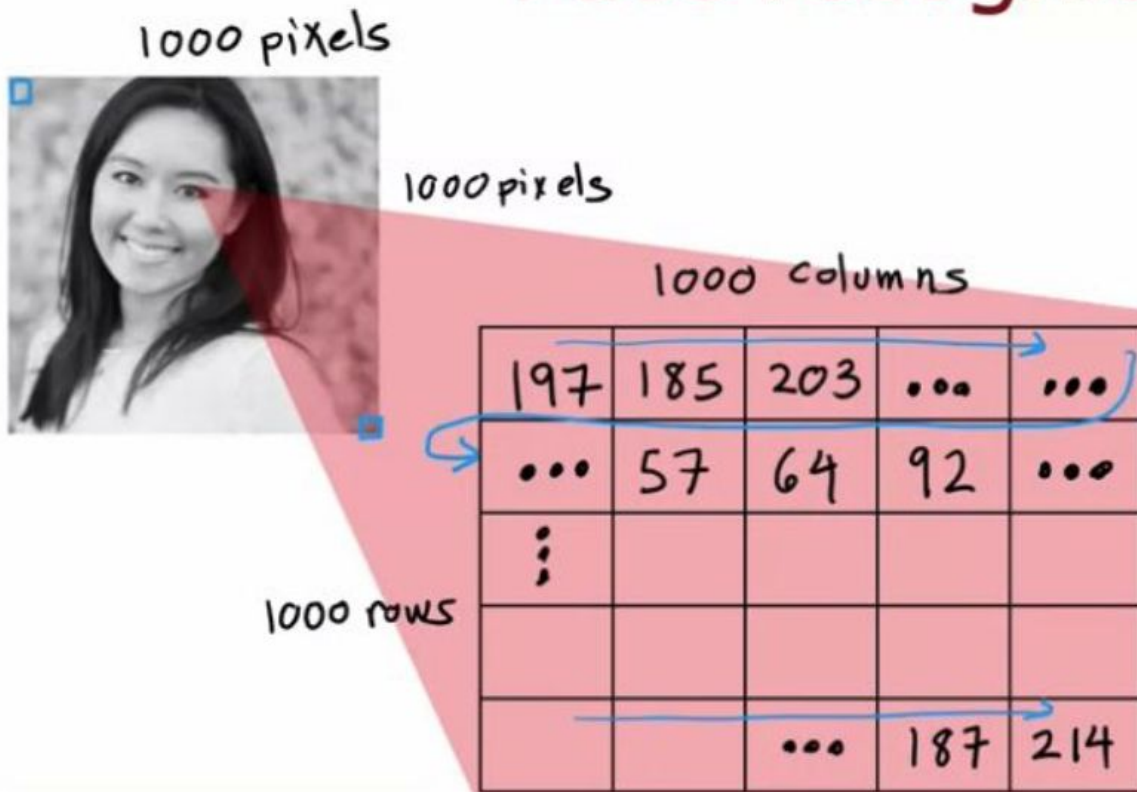
Aprendizado Supervisionado: Exemplo

- Reconhecimento de dígitos escritos a mão



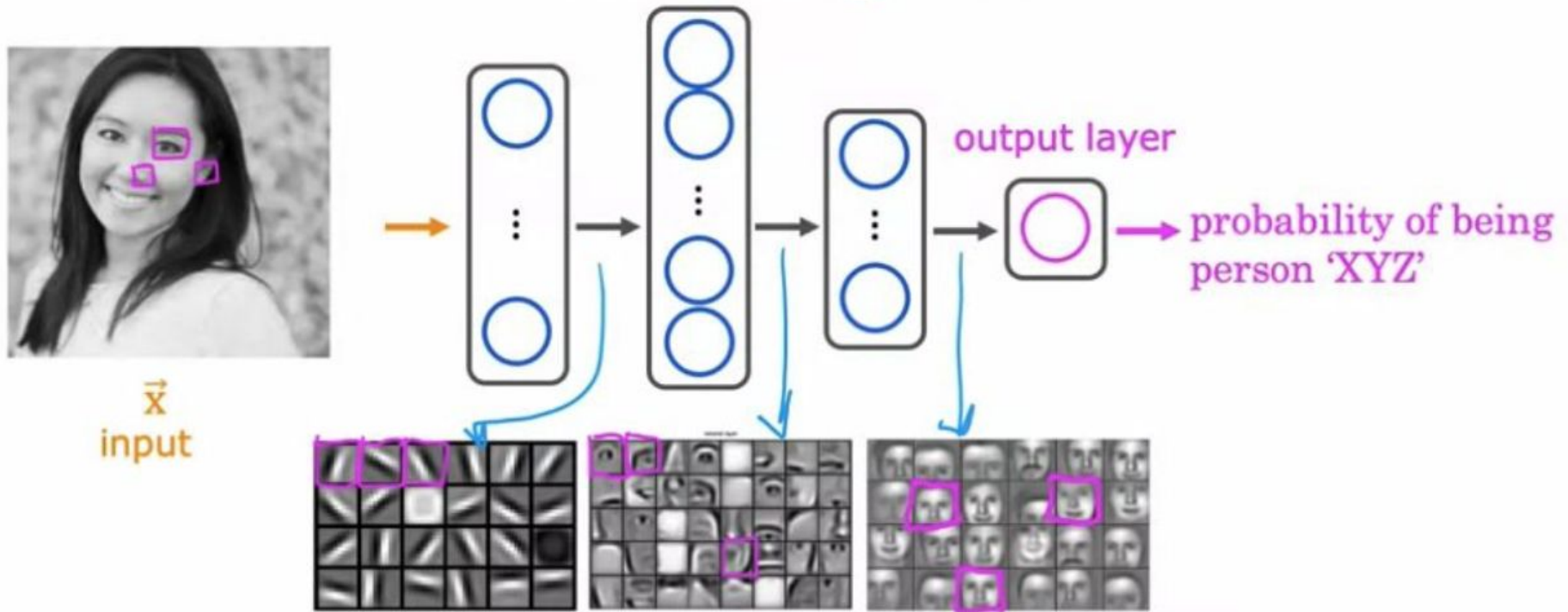
Aprendizado Supervisionado: Exemplo

Face recognition



Aprendizado Supervisionado: Exemplo

Face recognition



source: Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations
by Honglak Lee, Roger Grosse, Ranganath Andrew Y. Ng

Aprendizado Supervisionado: Exemplo

Classify



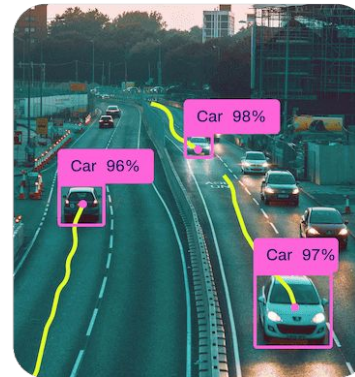
Detect



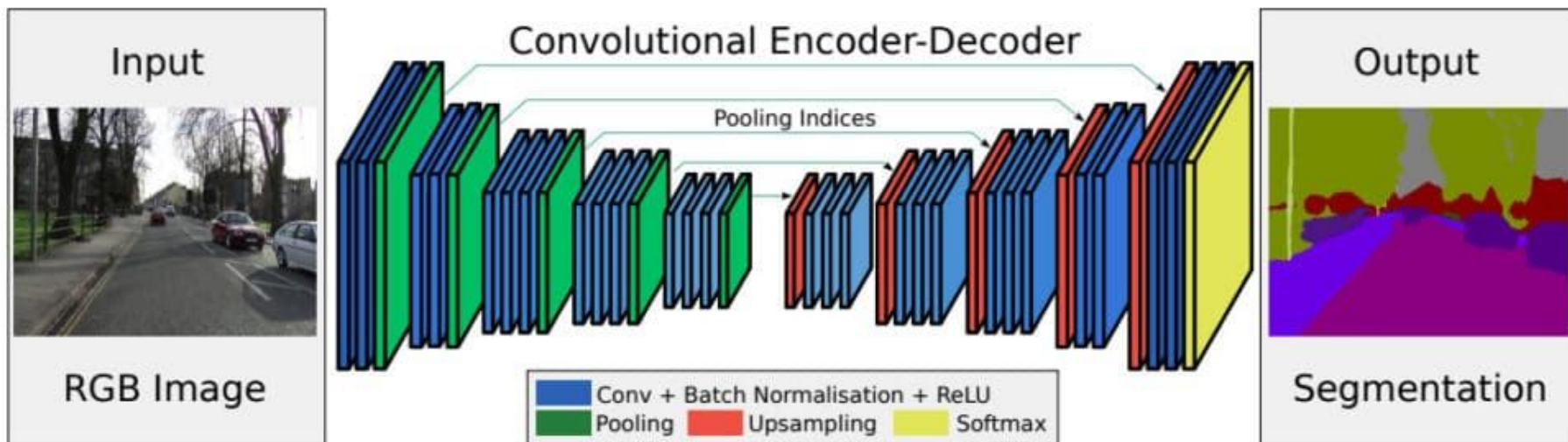
Segment



Track



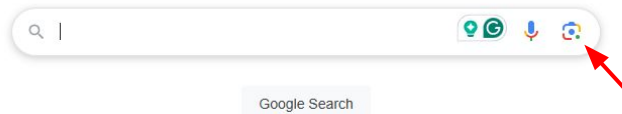
Pose



Machine Learning

- **Aprendizado Supervisionado:**
 - Dadas medições $(\mathbf{X}_1, Y_1), \dots, (\mathbf{X}_n, Y_n)$, aprender um modelo para **prever** Y_i baseado em \mathbf{X}_i
- **Aprendizado Não-Supervisionado:**
 - Dadas medições $(\mathbf{X}_1, \dots, \mathbf{X}_n)$, descobrir alguma **estrutura** com base em similaridade.

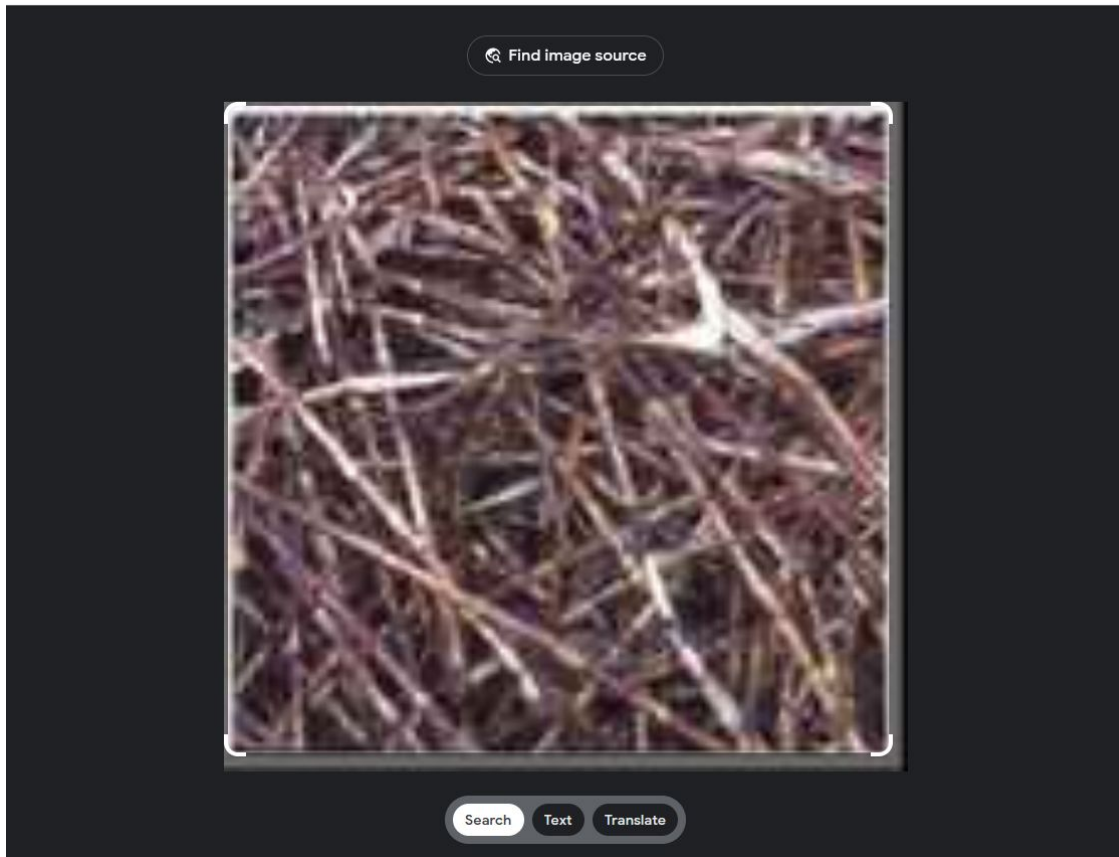
Aprendizado Não-Supervisionado: Exemplo



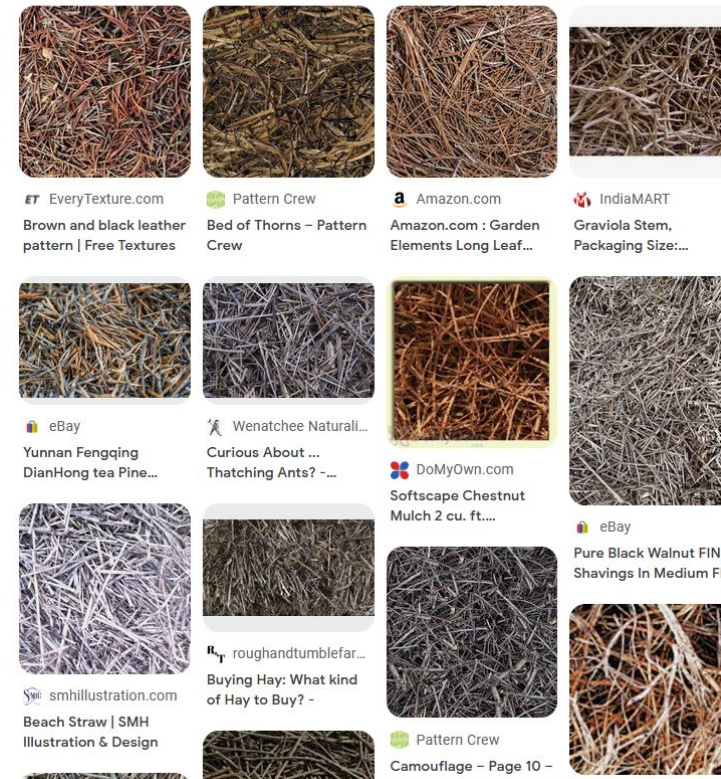
$X_i \rightarrow$ imagens na Internet

busca por estrutura (similaridade)

Google



Visual matches

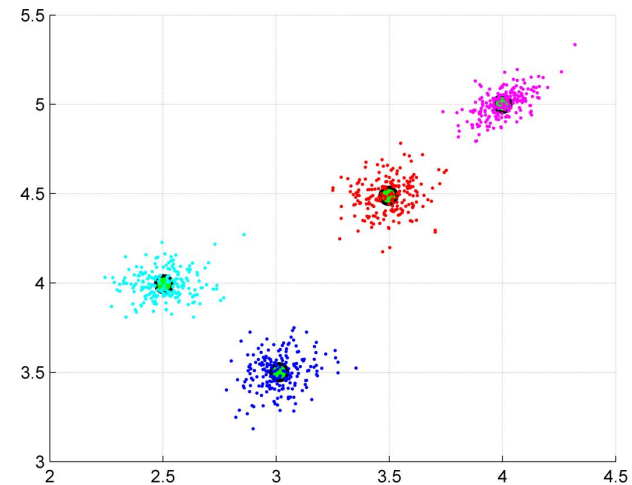
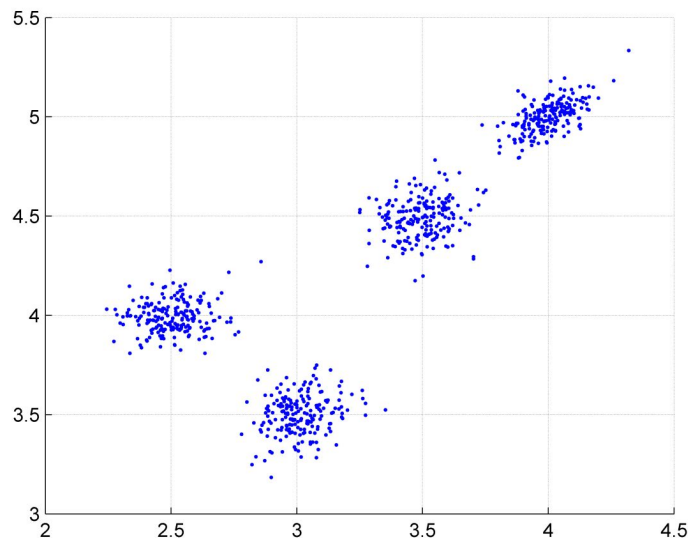


Aprendizado Não-Supervisionado: Exemplo

- *Market Segmentation:*

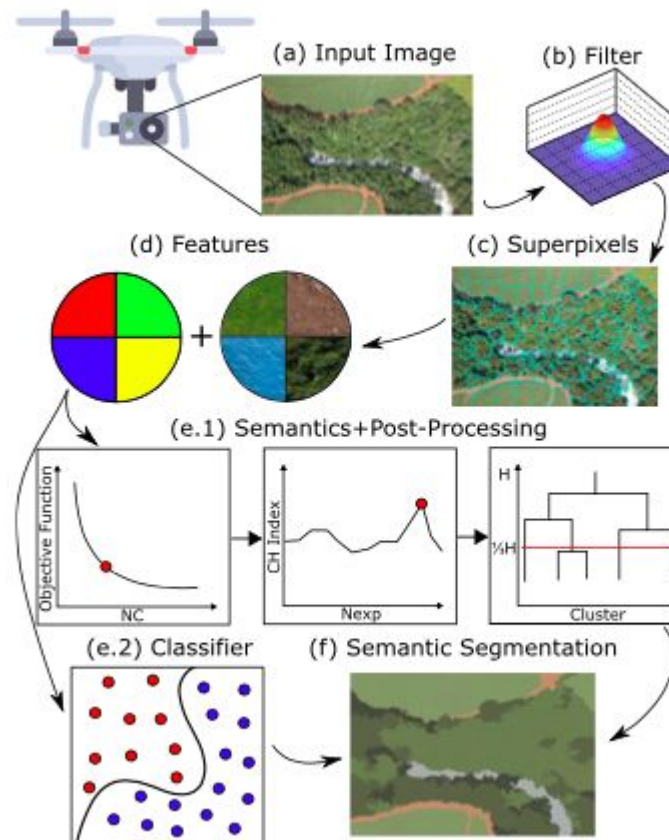
X_i = info de uma pessoa/cliente: renda, profissão, idade, distância do centro urbano mais próximo, etc.

- **objetivo:** identificar **subgrupos** de pessoas que podem ser mais receptivas a uma determinada forma de propaganda ou potenciais compradores de um produto particular.



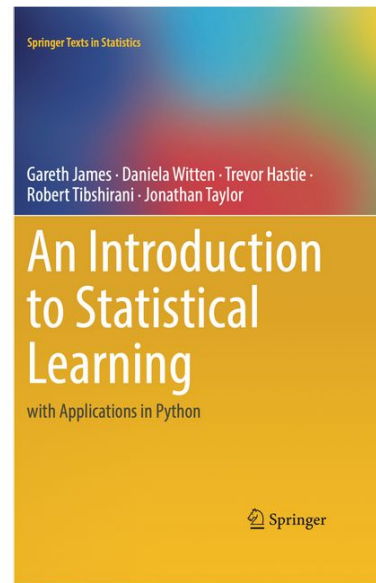
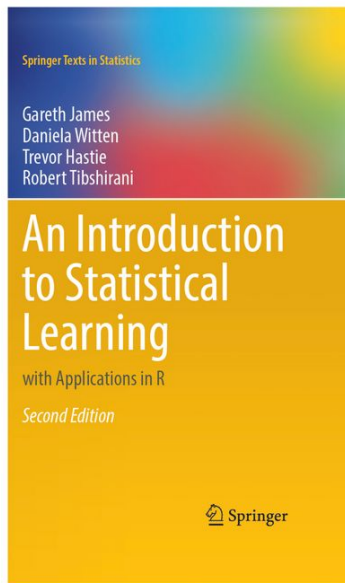
Aprendizado Não-Supervisionado: Exemplo

- Segmentação de Imagens por Regiões Homogêneas:



Fonte: Jaimes, B. R. A., Ferreira, J. P. K., & Castro, C. L. (2022). Unsupervised semantic segmentation of aerial images with application to uav localization. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.

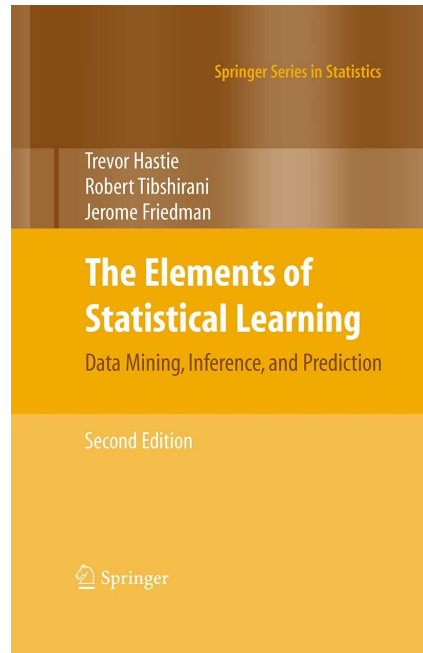
Material



James, Gareth, et al. *An introduction to statistical learning*. 3rd Edition. New York: Springer, 2022.

- Disponível para download gratuito em:
<https://www.statlearning.com/>

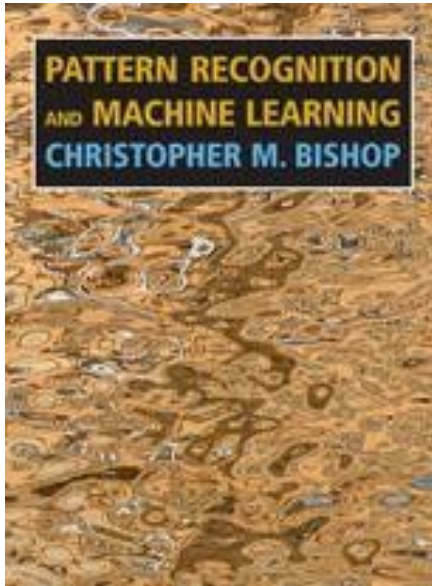
Material



Hastie, et al. *Elements of Statistical Learning*. 2nd Edition New York: Springer.

- Disponível para download gratuito em:
<https://hastie.su.domains/ElemStatLearn/>

Material

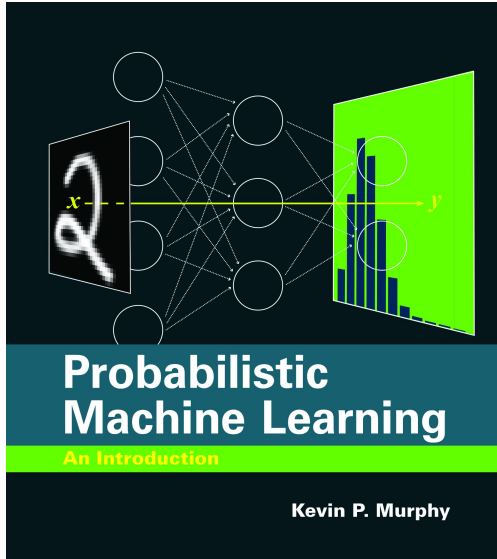


Christopher M. Pattern Recognition and Machine Learning. New York :Springer, 2006.

- Disponível em:

<https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>

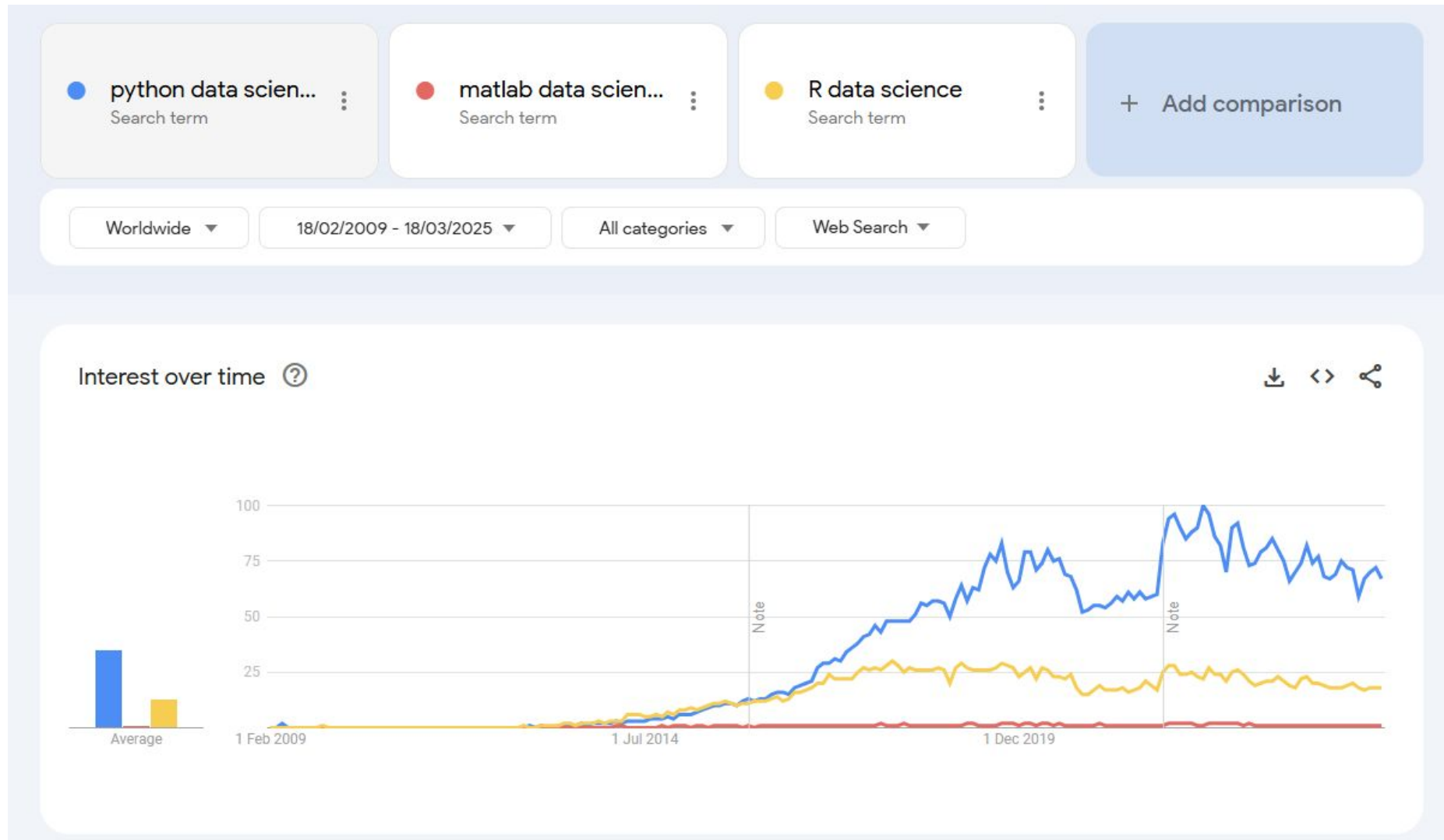
Material



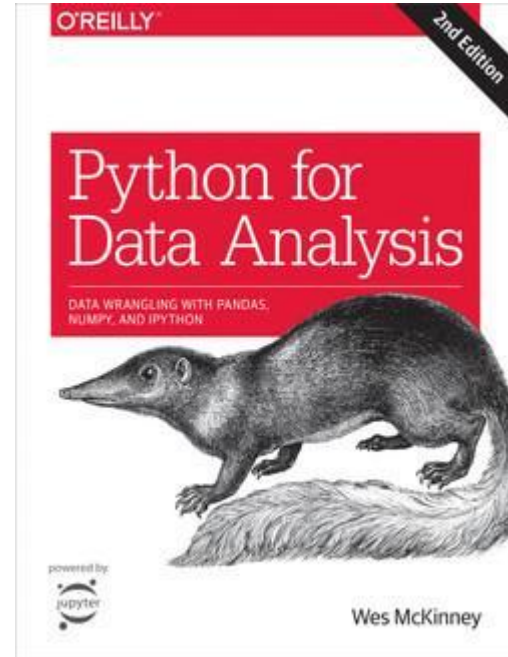
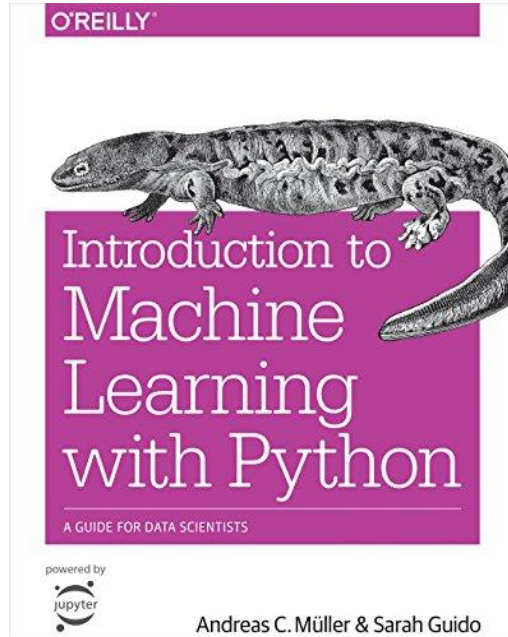
Kevin P. Murphy. Probabilistic Machine Learning. An Introduction. MIT Press, March 2022.

- Disponível em:
<https://probml.github.io/pml-book/book1.html>

Python vs R vs Matlab



Material - Python



- ver no Moodle:
 - python numpy-matplotlib tutorial
 - JAI-SBC sobre jupyter notebooks para ciência de dados

Plano de Aulas

Data	Assunto	Exercícios
19-03-2025	Apresentação do Curso	
26-03-2025	Fundamentos do Aprendizado de Máquina	Exercício 1
02-04-2025	Regressão Linear	
09-04-2025	Regressão Logística	Exercício 2
16-04-2025	Reamostragem - Validação Cruzada e Bootstrap	
23-04-2025	Seleção de Modelos e Regularização	Exercício 3
30-04-2025	-----	
07-05-2025	Prova	
14-05-2025	Métodos Baseados em Árvores	
21-05-2025	Máquinas de Vetores de Suporte	Exercício 4
28-05-2025	Explicação do TP Final	
04-06-2025	A Definir	
11-06-2025	Apresentação Preliminar - TP Final	
18-06-2025	A Definir	
25-06-2025	Entrega do TP Final	
Avaliações:	Exercícios Práticos: 40%	
	Prova: 30%	
	Trabalho Final: 30%	

Exemplos de Problemas

- Diagnóstico de Doenças Coronárias:

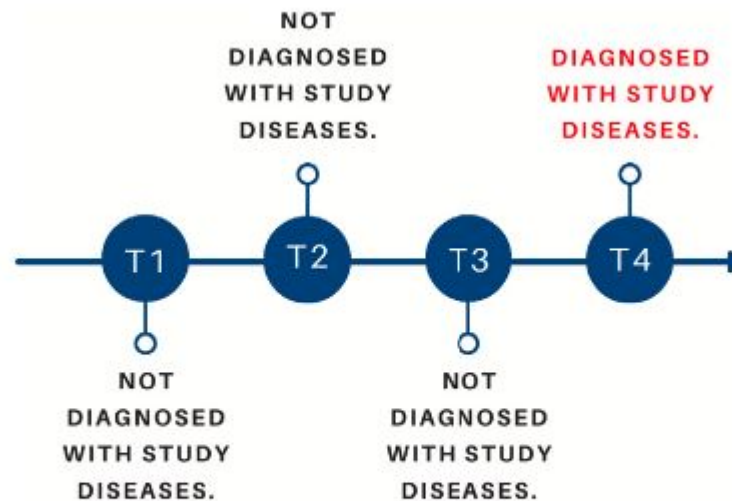


Fig. 2. Timeline illustrating the sequence of appointments and diagnoses of an arbitrary patient.

Fonte: Silva, C. A. O., Gonzalez-Otero, R., Bessani, M., Mendoza, L. O., & Castro, C. L. (2022). Interpretable risk models for sleep apnea and coronary diseases from structured and non-structured data. *Expert Systems with Applications*, 200.

PPGEE - Linha IC

- **Introdução à Inteligência Computacional**
- Técnicas Clássicas em Reconhecimento de Padrões
- Sistemas Nebulosos
- Redes Neurais Artificiais
- Sistemas Multiagentes
- Computação Evolucionária
- Otimização em Engenharia
- Otimização Multiobjetivo
- Análise e Detecção de Imagens
- Identificação de Sistemas
- Planejamento e Análise de Experimentos
- Algoritmos e Estruturas de Dados para Eng. Computação
- Aprendizado por Reforço
- etc.

Exemplos de Problemas

- Diagnóstico de Doenças Coronárias:
 - Dados Estruturados:
 - idade, sexo, altura, peso, pressão arterial, histórico de doenças diagnosticadas (ICD-10).
 - Dados Não-Estruturados:
 - campos textuais relativos ao comportamento do paciente e também a sua família
 - cirurgias, uso de tóxicos, doenças familiares

t al.

Expert Systems With Applications 200 (2022) 1

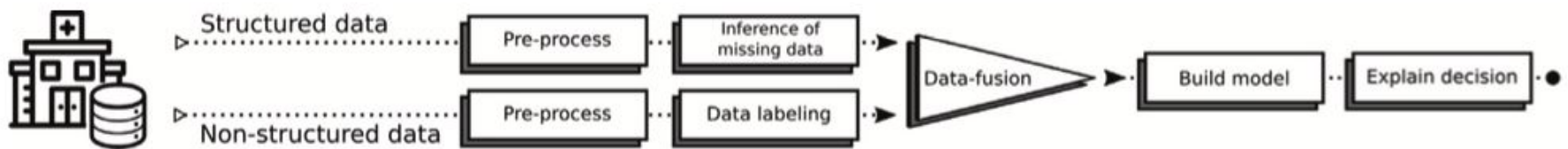


Fig. 1. Overview diagram of the proposed framework.

Exemplos de Problemas

- Diagnóstico de Doenças Coronárias:

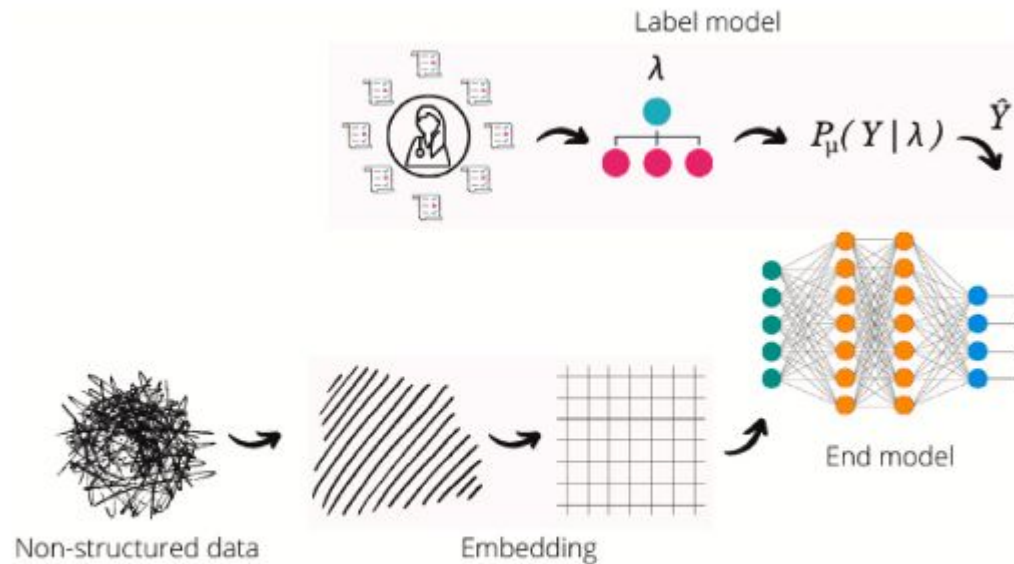


Fig. 3. Schema for generating relevant input variables from textual data.

Exemplos de Problemas

- Diagnóstico de Doenças Coronárias:

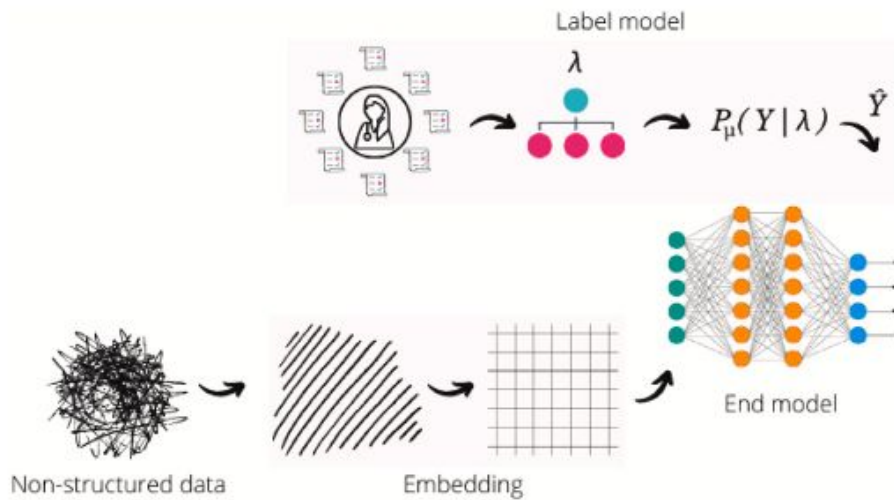


Fig. 3. Schema for generating relevant input variables from textual data.

Questions defined by the specialist and experienced with the *Weak Supervision*.

	ID	Questions
Surgeries	S52	Does the patient report a history of cardiac surgery?
	S53	Does the patient report a history of cardiovascular surgery?
	S54	Does the patient report a history of cancer surgery?
	S55	Does the patient report a history of lung surgery?
	S56	Does the patient report a history of brain trauma surgery?
	S74	Does the patient refer to bariatric surgery?
	S75	Does the patient refer to pancreatic surgery?
Families	F57	Does the patient report a history of a mother with a disease?
	F58	Does the patient have a history of a father with a disease?
	F59	Does the patient report a history of siblings with any illness?
	F60	Does the patient report a history of a grandmother with an illness?
	F61	Does the patient report a history of a grandfather with a disease?
	F62	Does the patient report a history of uncles with an illness?
	F76	Does the patient refer the mother with arterial hypertension?
	F77	Does the patient refer a father with hypertension?
	F78	Does the patient refer to the mother with acute myocardial infarction?
	F79	Does the patient refer a father with acute myocardial infarction?
	F80	Does the patient refer to the mother with atrial fibrillation?
	F81	Does the patient refer a father with atrial fibrillation?
	F82	Does the patient refer to the mother with sleep apnea?
	F83	Does the patient refer to a father with sleep apnea?
Toxic	T63	Does the patient report a history of drug use?
	T64	Does the patient report a smoking history?
	T65	Does the patient report a history of alcohol abuse?
	T66	Does the patient report a history of exposure to toxic substances?

Exemplos de Problemas

- Diagnóstico de Doenças Coronárias:

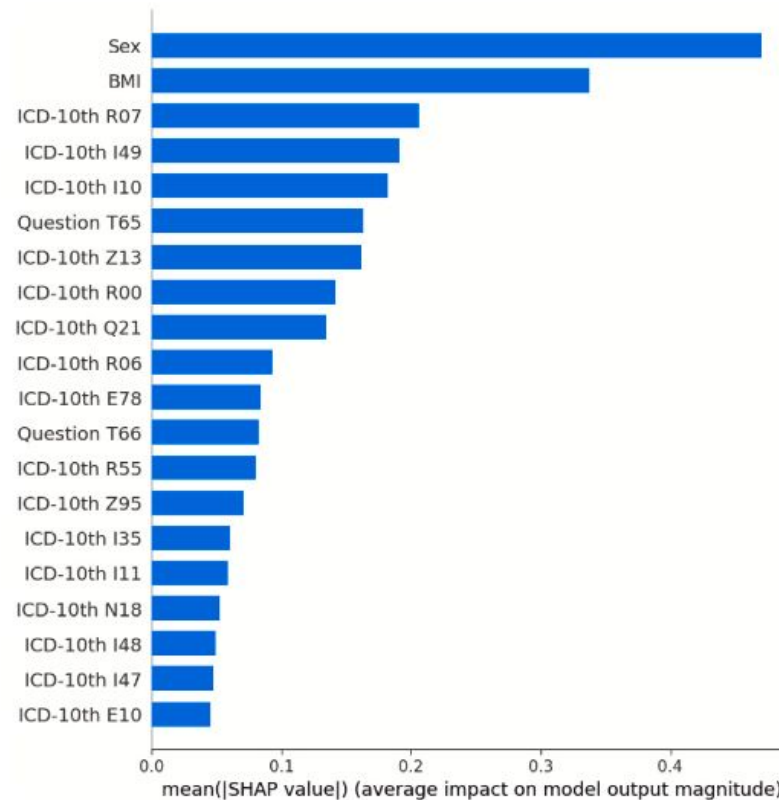


Fig. 4. SHAP values for the 20 most important features for CAD prediction using physical, medical history, and non-structured data. The higher the SHAP value of a variable, the greater is its importance to prediction.

Exemplos de Problemas

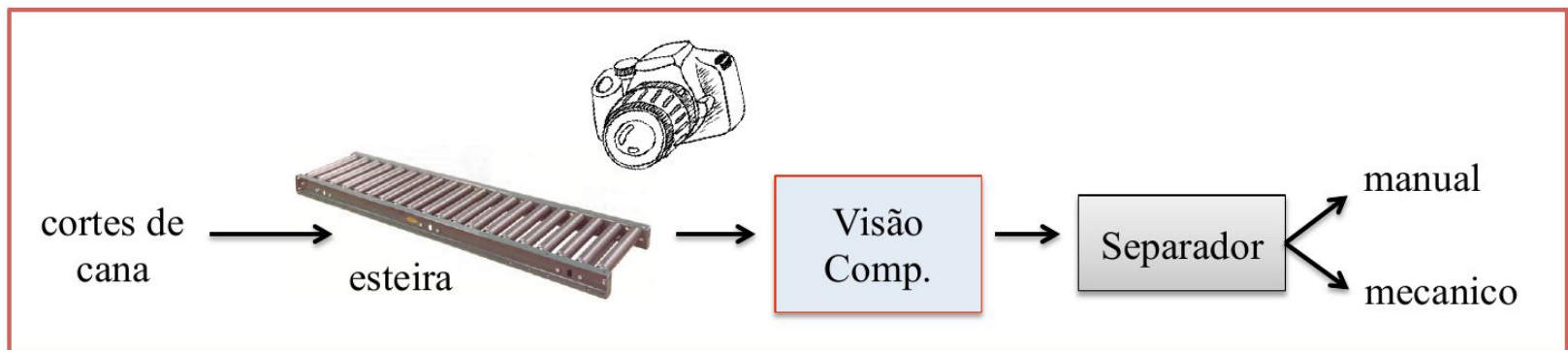
- Classificação automática de cortes de cana-de-açúcar:



Manual

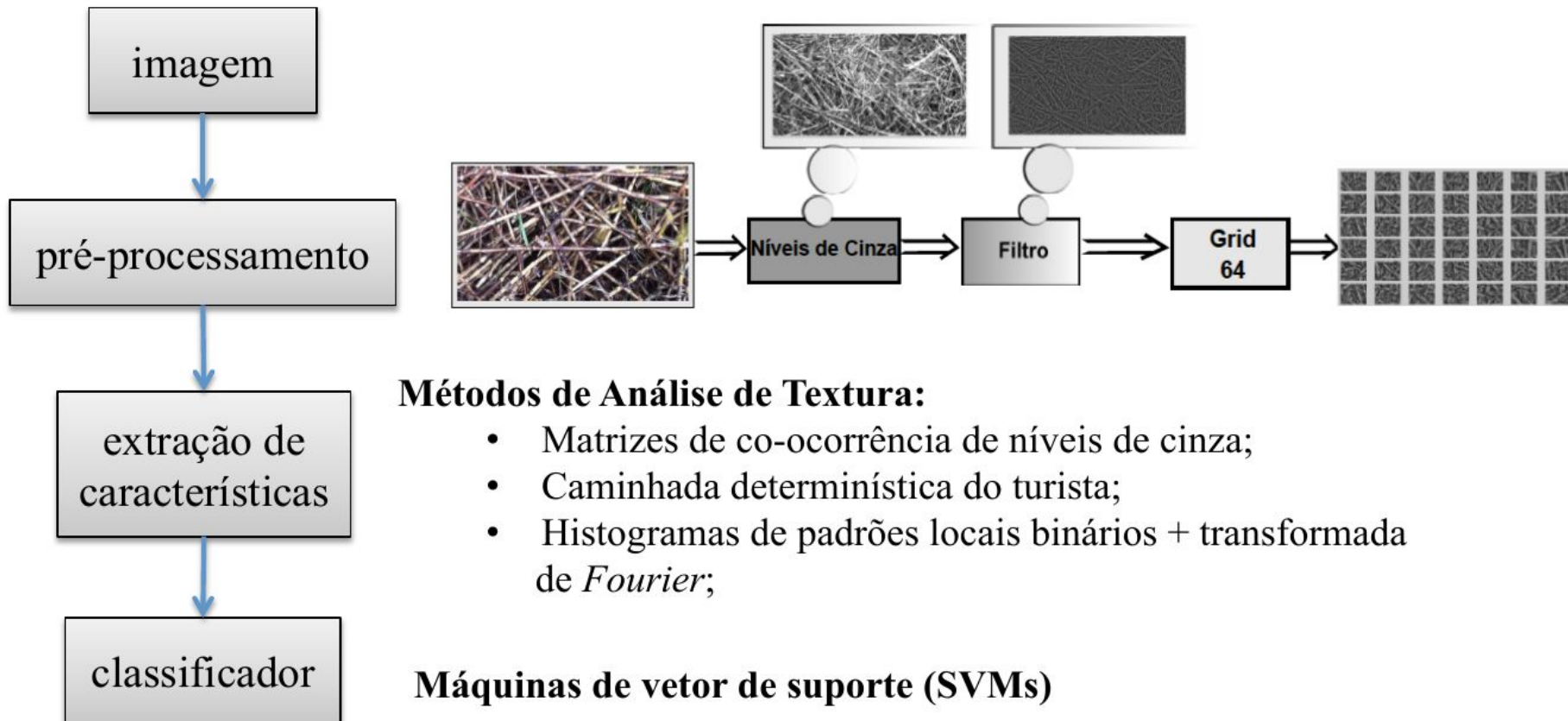


Mecânico



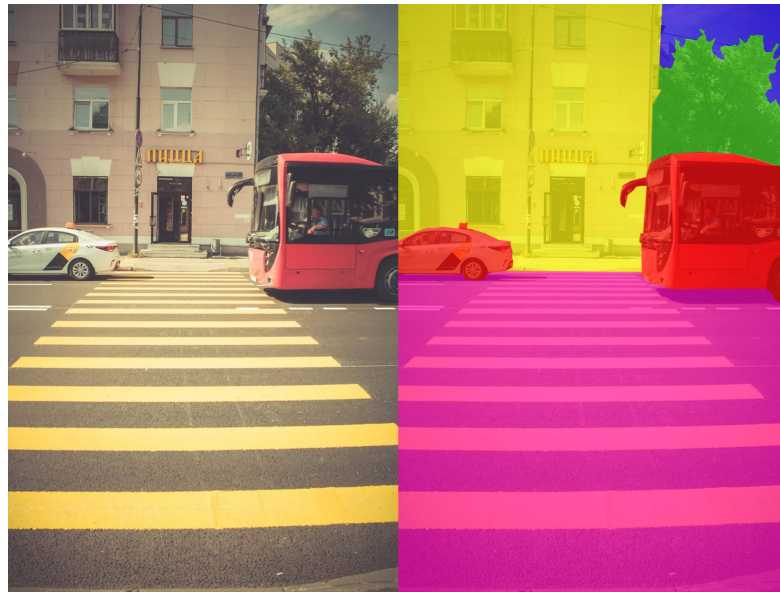
Exemplos de Problemas

- Classificação automática de cortes de cana-de-açúcar:



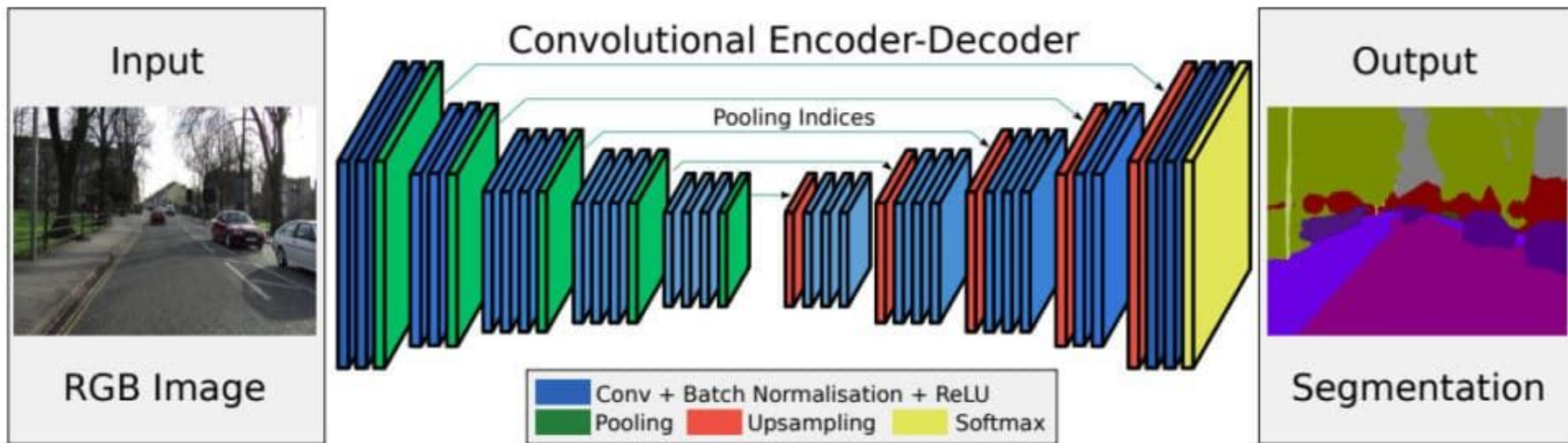
Exemplos de Problemas

- Segmentação Semântica de imagens



Exemplos de Problemas

- Segmentação Semântica de imagens



Exemplos de Problemas

- Geolocalização de VANTs por Imagens:



(a)

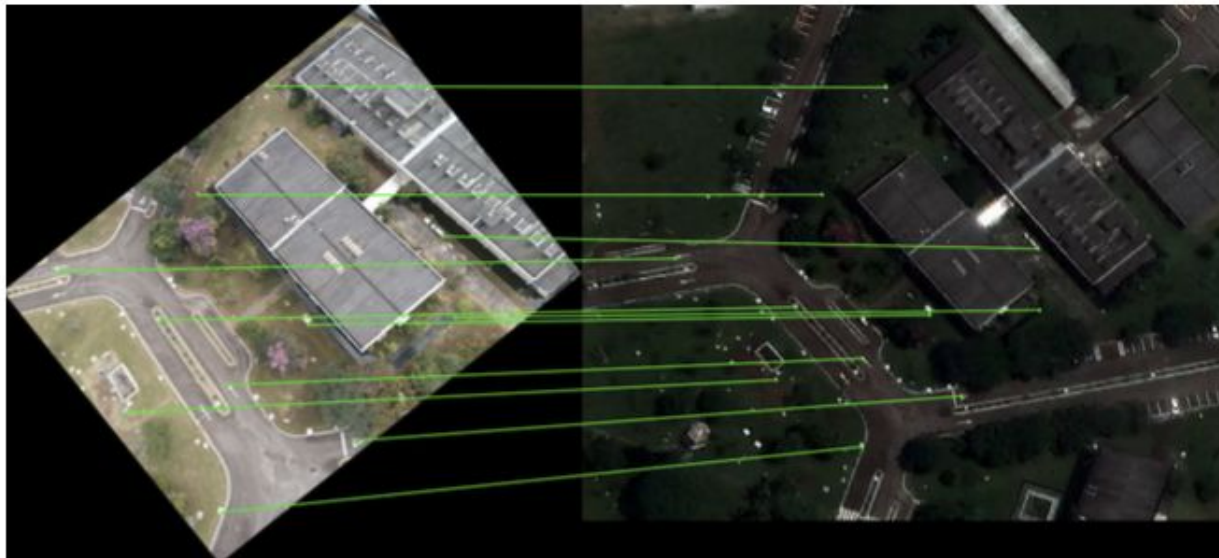


(b)

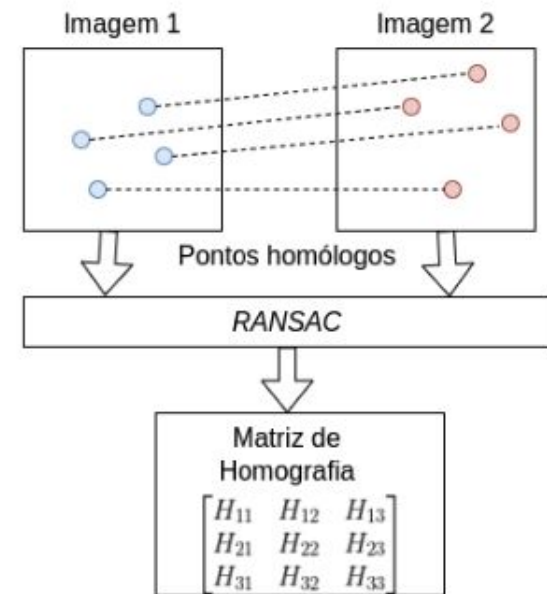
FIGURA 2.1 – (a) Imagem obtida por uma aeronave. (b) Imagem satelital obtida pelo Google Earth.

Exemplos de Problemas

- Geolocalização de VANTs por Imagens:
 - Técnica de Casamento de Pontos Característicos



(a)



(b)

Exemplos de Problemas

- Geolocalização de VANTs por Imagens:

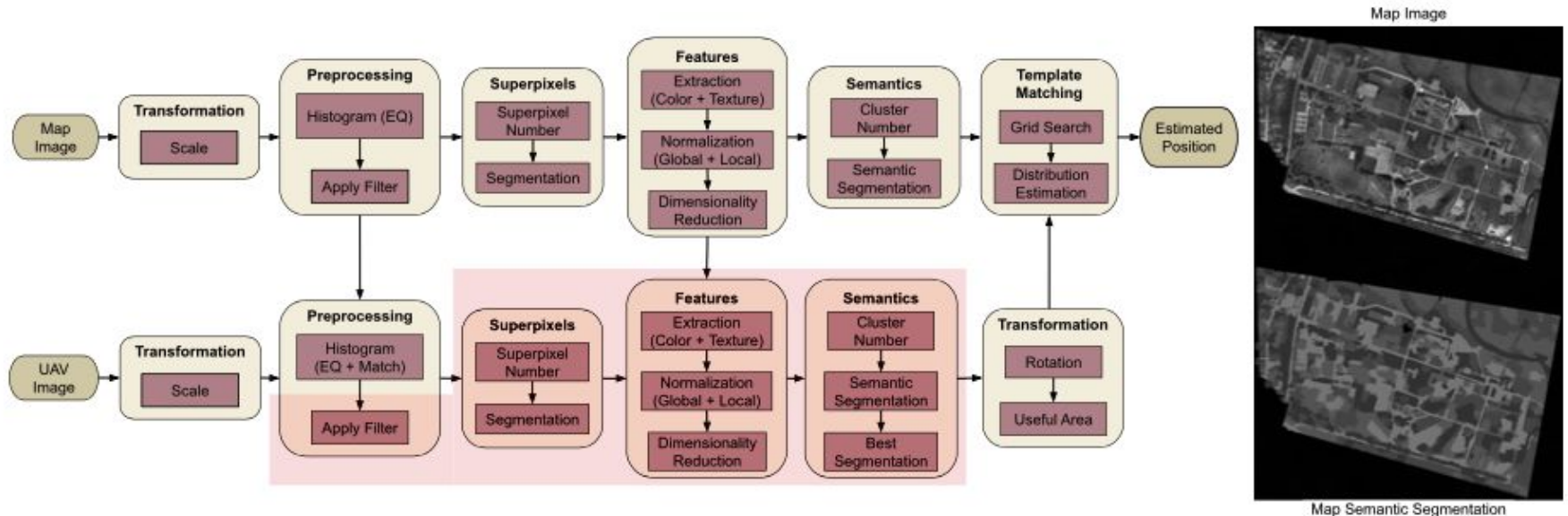


Fig. 2. (Left) Overview of methodology for UAV position estimation via template matching that includes, as a key stage, the proposed method of semantic segmentation which is highlighted in the red shaded area. (Right) Map image from SW dataset and its semantic segmentation.

Fonte: Jaimes, B. R. A., Ferreira, J. P. K., & Castro, C. L. (2022). Unsupervised semantic segmentation of aerial images with application to uav localization. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.

Exemplos de Problemas

- Geolocalização de VANTs por Imagens:



Fig. 3. Position estimators' comparison: black, original UAV route; yellow, route from [11]; red and blue, SegNet and DeepLab routes [13]; orange, ours.

Fonte: Jaimes, B. R. A., Ferreira, J. P. K., & Castro, C. L. (2022). Unsupervised semantic segmentation of aerial images with application to uav localization. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5.

Metodologia de Ensino

- Aulas semanais: **Terças-Feiras** - 13:30 às 16:30hs
 - Teoria: apresentação conteúdo teórico (~ 1:30 h) - presencial
 - Prática: exercícios computacionais (~ 1:30 h) - normalmente online (via MS Teams)
- Notas de aula, avisos, etc. (*MS Teams*)
- Os exercícios computacionais devem ser desenvolvidos na linguagem Python
- Trabalho Final com implementação em Python