

Curso Prático de Deep Learning - Aula 1

Aprendizado Visual Shallow

Hugo Oliveira¹, Roberto Cesar Jr.¹

¹Instituto de Matemática e Estatística (IME)
Universidade de São Paulo

16 de Janeiro, 2023



Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Grupo para Mensagens Rápidas (WhatsApp)

Creativision



Instrutor e Coordenador



Hugo Oliveira¹



Roberto Cesar Jr.²

¹<https://sites.google.com/view/oliveirahugo>

²<https://www.ime.usp.br/cesar/>

O que Veremos no Curso?

- Intuição sobre o que acontece durante o treino de uma rede neural
- Conceitos básicos de Deep Learning usando redes neurais artificiais
- Ferramentas, frameworks e ambientes de programação para desenvolvimento de soluções com Deep Learning

O que Não Veremos no Curso?

- Bases matemáticas aprofundadas do Deep Learning
- Métodos ainda não tão consolidados na literatura
- Abordagens aprofundadas fora do domínio de imagens
- Modelos demasiadamente pesados ou demorados para serem executados em 1 única GPU

Ementa

- Aula 1 (16/01) - Shallow Learning
 - Ambiente
 - Python & bibliotecas
 - Features “handcrafted” para imagens
 - Shallow Learning
 - Bag of Visual Words
- Aula 2 (18/01) - Introdução a Redes Neurais e MultiLayer Perceptrons
 - Fundamentos de Redes Neurais
 - MultiLayer Perceptrons (MLPs)
 - Classificação de imagens usando MLPs

Ementa do Curso

- Aula 3 (20/01) - Redes Neurais Convolucionais, Segmentação e Detecção
 - Redes Neurais Convolucionais (CNNs)
 - Segmentação
 - Detecção de Objetos e Segmentação de Instância
- Aula 4 (23/01) - Modelagem Generativa
 - Modelagem Generativa
 - AutoEncoders
 - Generative Adversarial Networks (GANs)

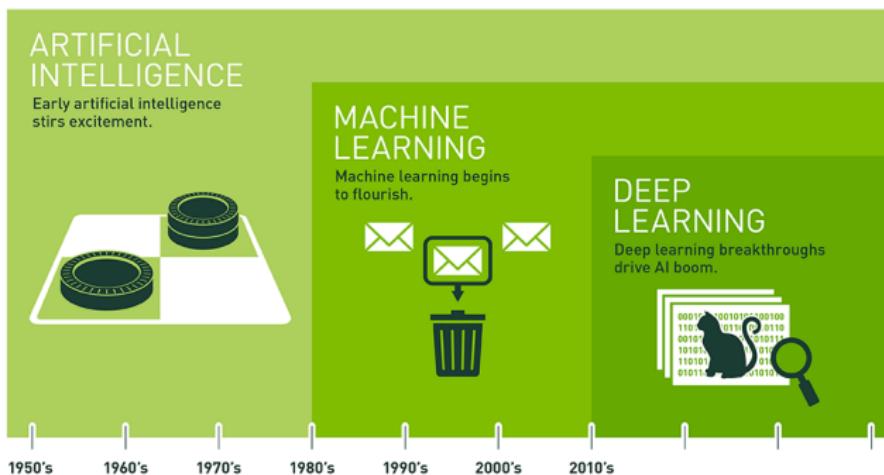
Ementa do Curso

- Aula 5 (27/01) - Few-Shot Learning
 - Desafios da aprendizagem em cenários Low-Shot
 - Domain Adaptation e Transfer Learning
 - Representation Learning
 - Meta-Learning
- Aula 6 (30/01) - Métodos Modernos
 - Tradução de Imagens
 - Open Set Recognition
 - Módulos de Attention
 - Transformers
 - Síntese Text-to-Image
 - Redes Forward-Forward
 - gcForests
 - ...

Organização do Curso

- Aulas teóricas e práticas
- Teórica: slides
- Prática (Colab):
 - Demos (Extras) – a serem executados como atividade extra, embora altamente recomendados
 - Demos – demonstração de conceitos sendo executados em tempo real nos jupyter notebooks
 - Práticas – atividade para realização em sala de aula sob supervisão do professor
 - Challenges – atividades práticas para realização em horário extra-classe que irão demandar mais esforço de implementação dos alunos

Introdução a Deep Learning



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.

Figura: Ilustrando a relação entre Deep Learning, Aprendizado de Máquina e Inteligência Artificial [1]

O que é Aprendizado de Máquina?

Definição Abstrata¹

Um algoritmo capaz de aprender com a experiência, nos permitindo lidar com tarefas muito difíceis para serem solucionadas com algoritmos explicitamente programados por seres humanos.

Definição Prática

Um algoritmo parametrizado, capaz de encontrar o conjunto de parâmetros que melhor se aproxima do comportamento desejado.

¹ Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, 2016. <http://www.deeplearningbook.org> 11/82

O que motivou Deep Learning?

- O mundo real não é estruturado
- Dados do mundo real possuem uma alta variabilidade
- Um desafio para a Inteligência Artificial é aprender representações e modelos capazes de lidar com tais adversidades

O que motivou Deep Learning?

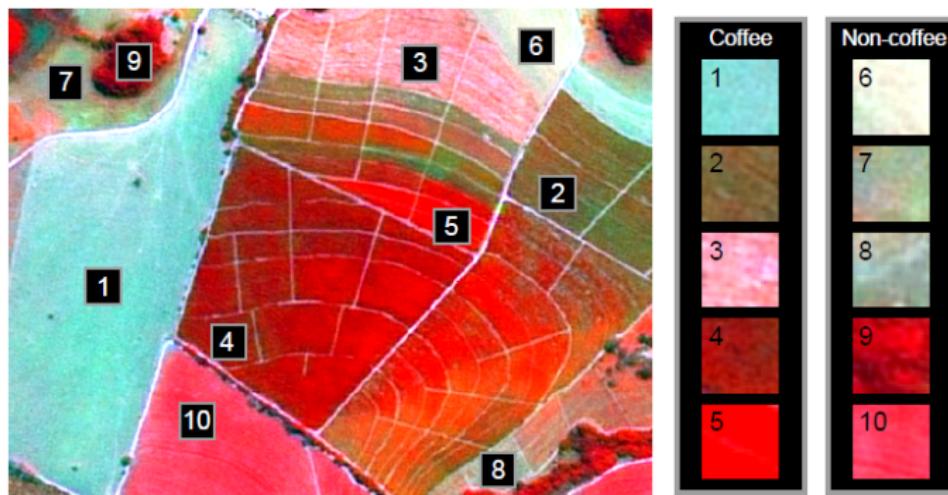


Figura: Exemplo de aplicação onde o dado possui diferentes variações.

O que motivou Deep Learning?

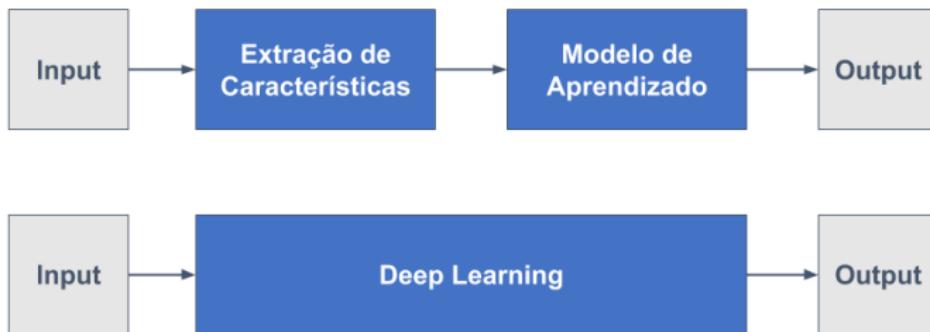
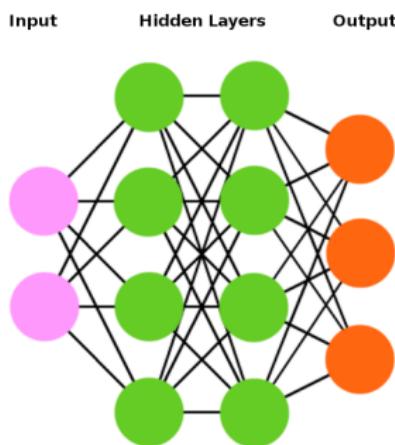


Figura: Diferença entre Aprendizado de Máquina tradicional e Deep Learning.

O que é Deep Learning?

- Tipo de Aprendizado de Máquina
- Baseado em Redes Neurais profundas
 - Mais de uma camada escondida
- Vasta gama de aplicações e **variações do modelo neural**

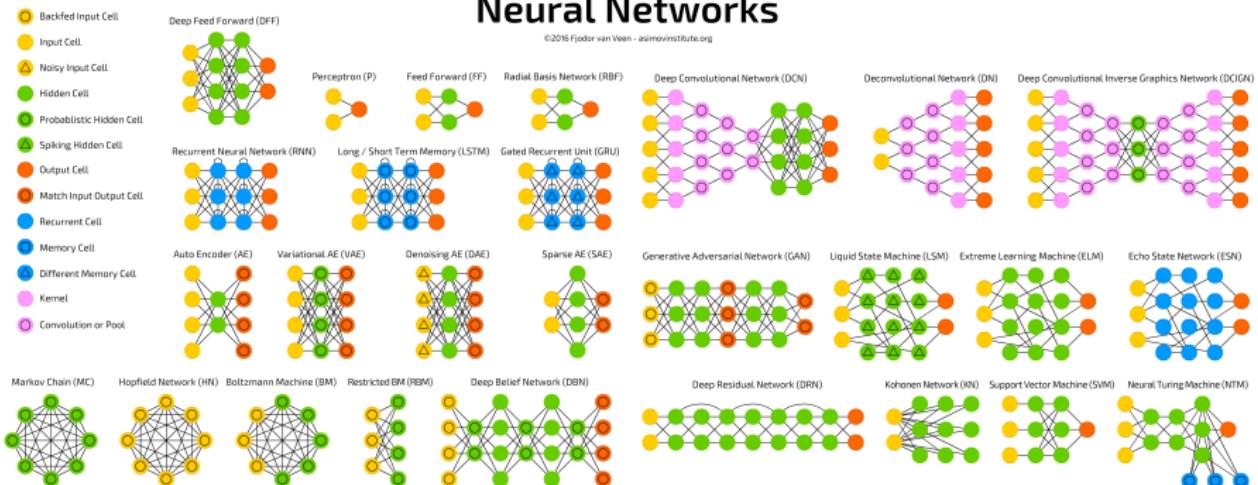


“Zoológico” de Arquiteturas

Creativision

A mostly complete chart of
Neural Networks

©2016 Fjodor van Veen - ai-movinstitute.org



O que é Deep Learning?

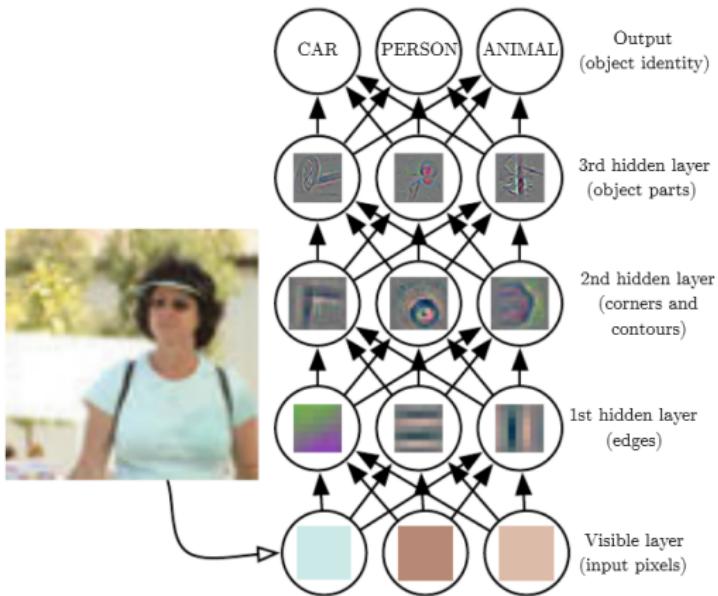


Figura: Exemplo [2] mostrando como Deep Learning parte de representações simples para mais complexas

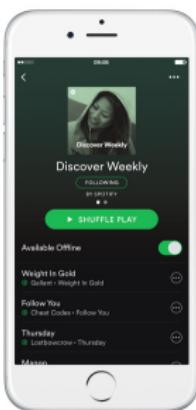
Aplicações

Creativision

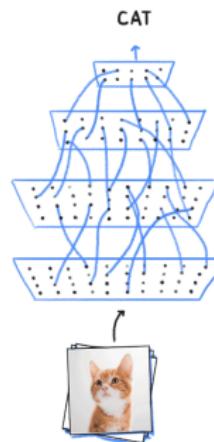
- Transferência de Estilo



- Recomendação de conteúdo



- Classificação



Aplicações

Creativision

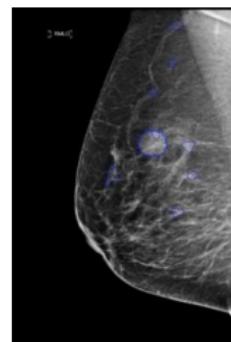
- Carros Autônomos



- Vigilância



- Imagens Médicas



O que é Deep Learning

Importância do Deep Learning

É difícil compreender a importância do avanço tecnológico que representa a chegada do Deep Learning sem explicar como funciona Shallow Learning

Ambiente Computacional

- Python
 - Linguagem e biblioteca padrão
- Pytorch
 - Framework de processamento vetorial e Deep Learning
- Google Colab
 - Serviço de Cloud Gratuito da Google
 - Ambiente de programação, plots e anotações

Google Colaboratory¹



- Construído com base no Jupyter
- Serviço de Cloud *gratuito* para desenvolvimento de I.A.
- 1 GPU disponível por até 12 horas contínuas
- Permite diferentes frameworks (Keras, Tensorflow, Pytorch)

¹ <https://colab.research.google.com/>

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Aprendizado Visual Clássico



Reconhecimento Visual Shallow

Desde o surgimento de algoritmos de aprendizado shallow modernos (i.e. SVM, Random Forest, AdaBoost, etc) nos anos 1990, a automação de tarefas de reconhecimento visual começou a ser possível para tarefas não triviais.

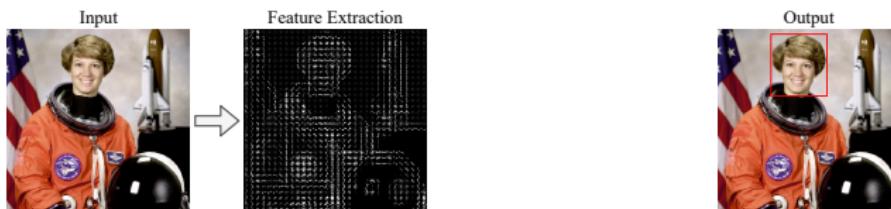
Aprendizado Visual Clássico



Reconhecimento Visual Shallow

Os primeiros algoritmos para a solução de problemas clássicos de Visão Computacional (i.e. classificação de cenas, segmentação de objetos, detecção de objetos) permitiram que algumas tarefas mais simples do mundo real fossem automatizadas com sucesso. Casos notáveis incluem reconhecimento de faces, tracking de seres humanos em vídeo e OCR.

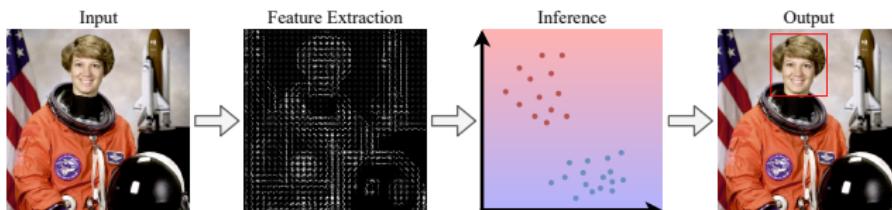
Aprendizado Visual Clássico



Reconhecimento Visual Shallow

O primeiro passo para atingir reconhecimento visual por meio de algoritmos de aprendizado shallow é selecionar extratores de características relevantes para os dados e a tarefa que se quer resolver. Os *features* obtidos por esses extratores podem codificar informação de forma (i.e. **Histogram of Oriented Gradients**), textura (i.e. **Gray Level Co-occurrence Matrix**), cores (i.e. **Global Color Histogram**) ou frequências (i.e. **Wavelets**).

Aprendizado Visual Clássico



Reconhecimento Visual Shallow

Depois da extração de características, um algoritmo de aprendizado (i.e. classificador, regressor, clusterizador, redutor de dimensionalidade, etc) pode ser treinado usando essas características extraídas da imagem para se ajustar aos dados/rótulos desejados para realizar a predição necessária à tarefa.

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Extração de Features

- Antes de Deep Learning, extração de features era necessária para lidar com dados não estruturados
- Features boas são desenhadas visando realçar e compactar as propriedades importantes do objeto
 - É difícil saber quais propriedades são importantes para o domínio e a tarefa em mãos
- Diferentes domínios possuem diferentes tipos de features

Extração de Features

- Alguns domínios possuem features naturais
 - Informações de prontuário médico (i.e. idade, sexo, peso, presença/ausência de sintomas etc)
 - Mercado imobiliário (i.e. bairro, andar, número de quartos, número de banheiros etc)
- Outros domínios precisam de algoritmos de extração de features
 - Imagens (i.e. forma, textura, cor etc)
 - Sinais Temporais (i.e. densidade espectral, frequências mais importantes, amplitude média etc)

Tipos de Features

- Features Categóricas

- Possui divisão bem definida entre categorias (i.e. bairro – Ouro Preto/Liberdade/Savassi/Lourdes, tipo – apartamento/casa/kitnet, cozinha – s/n, cor – branca/amarela/azul/verde, piscina – s/n etc)

- Features Numéricas

- Valores contínuos ou discretos que possuem uma relação numérica com valores próximos (i.e. andar, número de quartos, metros quadrados, número de banheiros etc)

Extração de Features em Imagens

- Imagens possuem muita redundância entre pixels vizinhos
- Como se livrar da informação redundante e extrair apenas as características mais importantes?

Extração de Features em Imagens

- Algumas categorias principais de features em imagens
 - Bordas
 - Frequência
 - Cor
 - Textura
 - Forma
- Diferentes domínios da área de Visão Computacional usam diferentes combinações desses features
- Dependendo da tarefa, os features podem ser extraídos da imagem toda (features globais) ou de patches (features locais)

Extração de Features em Imagens

- Bordas
 - Canny
 - Prewitt
 - Laplace

Extração de Features em Imagens

- Frequências
 - Wavelets
 - Discrete Fourier Transform (DFT)
 - Discrete Cosine Transform (DCT)

Extração de Features em Imagens

- Features de Cor

- Global Color Histogram (GCH)
- Border/Interior Pixel Classification (BIC)
- Color-Based Clustering (CBC)

Extração de Features em Imagens

- Features de Textura

- Gray-Level Co-occurrence Matrix (GLCM)
- Local Binary Patterns (LBP)
- Wavelets

Extração de Features em Imagens

- Features de Forma

- Histogram of Oriented Gradients (HOG)
- Scale-Invariant Feature Transform (SIFT)
- Speeded Up Robust Features (SURF)
- Extended Morphological Profiles (EMP)

Features em Imagens

Demo (Extra) - Features de Imagens

image_features.ipynb

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Pipeline do Aprendizado de Máquina

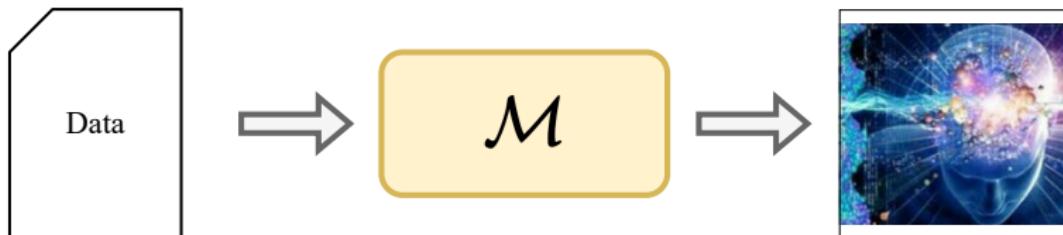


Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

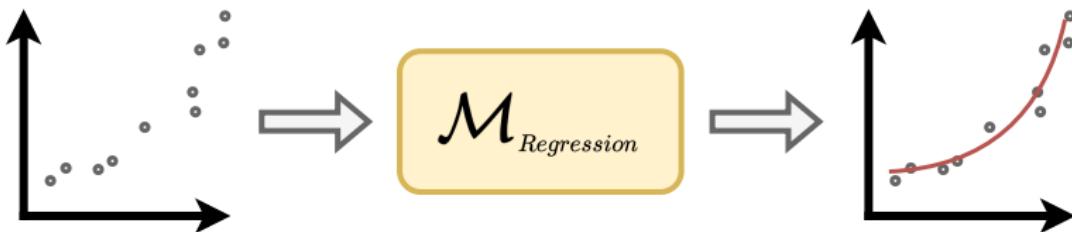


Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina



Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

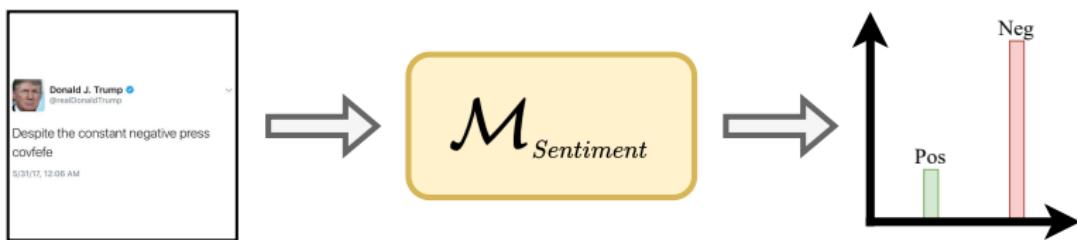


Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

Creativision



Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

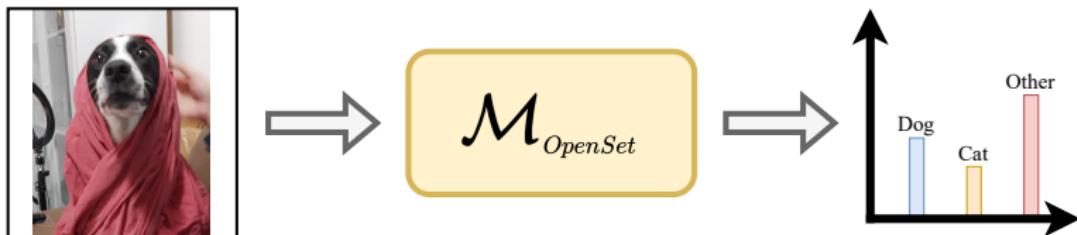


Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

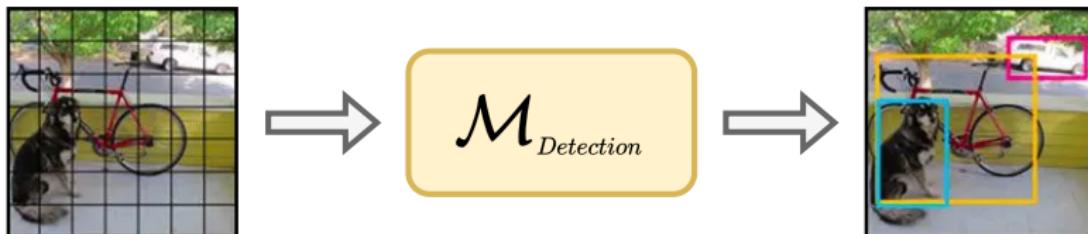
Creativsion

Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

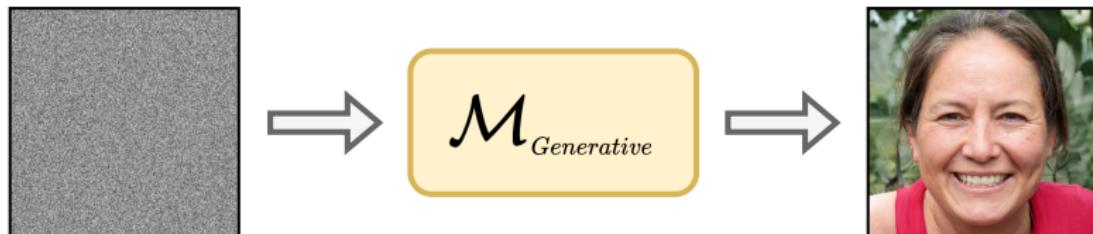


Figura: Pipeline de uma abordagem de Machine Learning.

Pipeline do Aprendizado de Máquina

Creativision



Figura: Pipeline de uma abordagem de Machine Learning.

Shallow Learning em Imagens

- Abordagem clássica para visão computacional
- Algoritmos de extração de características desenhados “manualmente”
- Modelos estatísticos para o reconhecimento automático de padrões

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Mapa de Métodos

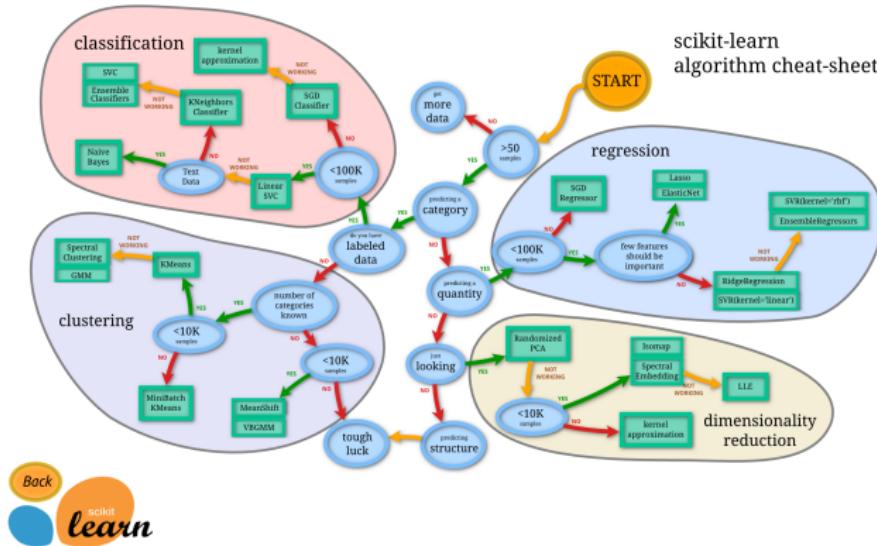


Figura: Mapa do scikit-learn¹ indicando os métodos indicados para cada tipo de problema.

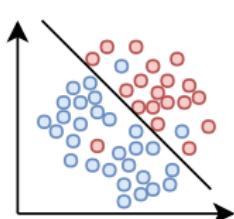
¹ http://scikit-learn.org/stable/tutorial/machine_learning_map/index.html

Tipos de Tarefas

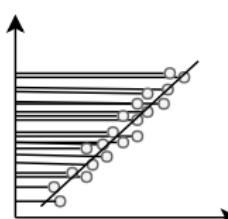
- Classificação
 - Predizem classes
- Regressão
 - Predizem valores numéricos
- Clusterização
 - Acham os clusters naturais presentes no conjunto de dados
- Redução de Dimensionalidade
 - Diminuem a dimensionalidade de dados para propósitos de análise ou visualização por humanos

Tipos de Tarefas

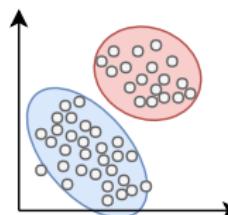
Creativision



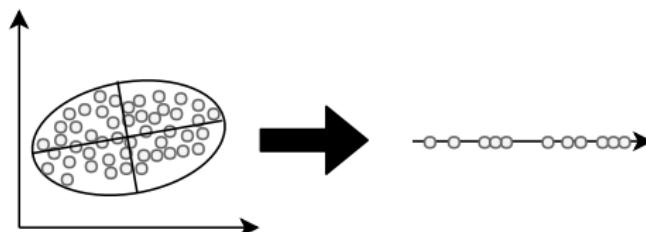
(a) Classificação



(b) Regressão



(c) Clusterização



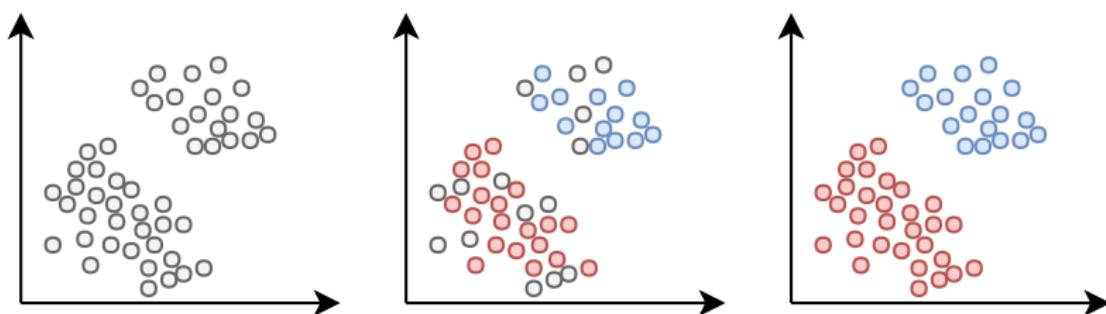
(d) Redução de Dimensionalidade

Figura: Algoritmos de acordo com os tipos de tarefas.

Tipos de Rotulação

- Algoritmos Não-Supervisionados
 - Nenhum rótulo
- Algoritmos Semi-Supervisionados
 - Amostras rotuladas e não rotuladas
- Algoritmos Supervisionados
 - Amostras rotuladas

Tipos de Rotulação



(a) Não-Supervisionado

(b) Semi-Supervisionado

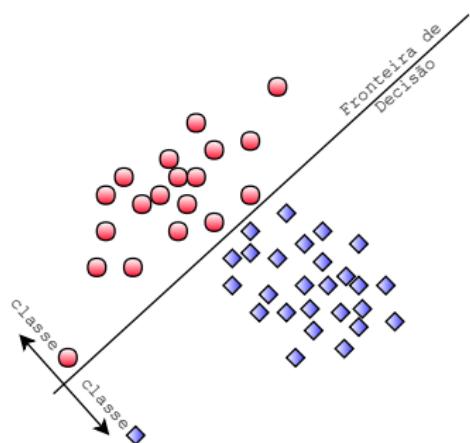
(c) Supervisionado

Figura: Algoritmos de acordo com os tipos de rotulação.

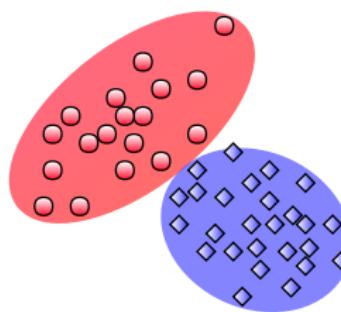
Tipos de Modelagem

- Modelos Generativos
 - Modelam a distribuição que gerou os dados
- Modelos Discriminativos
 - Modelam fronteiras de decisão entre diferentes classes

Tipos de Modelagem



(a) Modelo Discriminativo



(b) Modelo Generativo

Figura: Tipos de Modelagem.

Clusterização

- Aprendizado não-supervisionado com o objetivo de achar aglomerados naturais nas distribuições dos dados
- Só necessita dos features relacionados aos dados e de alguns hiperparâmetros (bandwidth, numero de aglomerados esperado, etc)
- Otimização de um *modelo generativo* que delineie as regiões no espaço de características em que as amostras estão mais concentradas

Algoritmos de Clusterização

- k -Means
- Gaussian Mixtures
- MeanShift
- BIRCH

Regressão

- Aprendizado supervisionado que visa predizer um valor numérico de uma variável aleatória com base em outras variáveis aleatórias
- Requer medições dos valores da função para as variáveis em múltiplos pontos
- Otimização de uma função que correlaciona duas ou mais variáveis

Algoritmos de Regressão

- Linear Regression
- Logistic Regression
- Ridge Regression
- Support Vector Regression

Classificação

- Separação de amostras em classes
- Requer dados rotulados
- Otimização de uma fronteira de decisão

Classificação

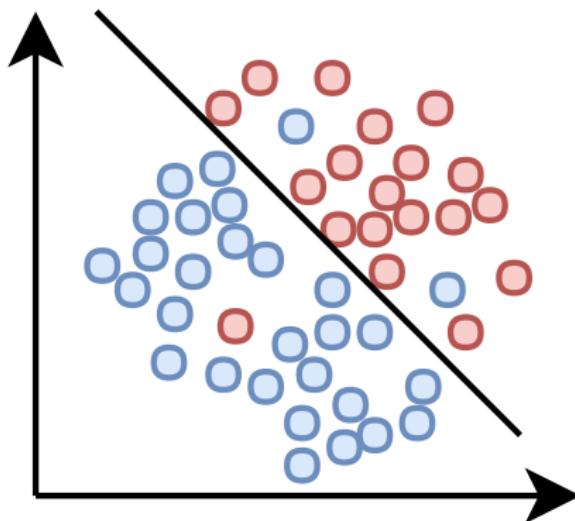
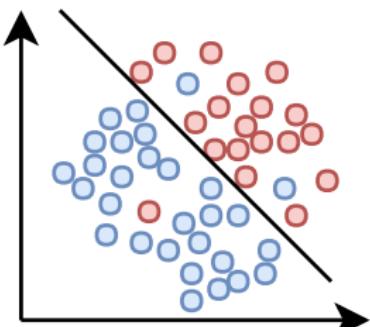
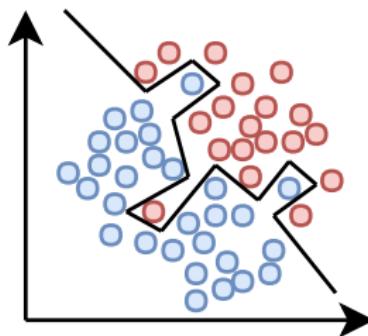


Figura: Tarefa de Classificação.

Overfitting em um Classificador



(a)



(b)

Figura: Fit razoável (a) e overfit (b) de uma fronteira de decisão em um classificador.

Algoritmos de Classificação

- k -Nearest Neighbors (kNN)
- Decision Trees (DT)
- Support Vector Machine (SVM)

k-Nearest Neighbors

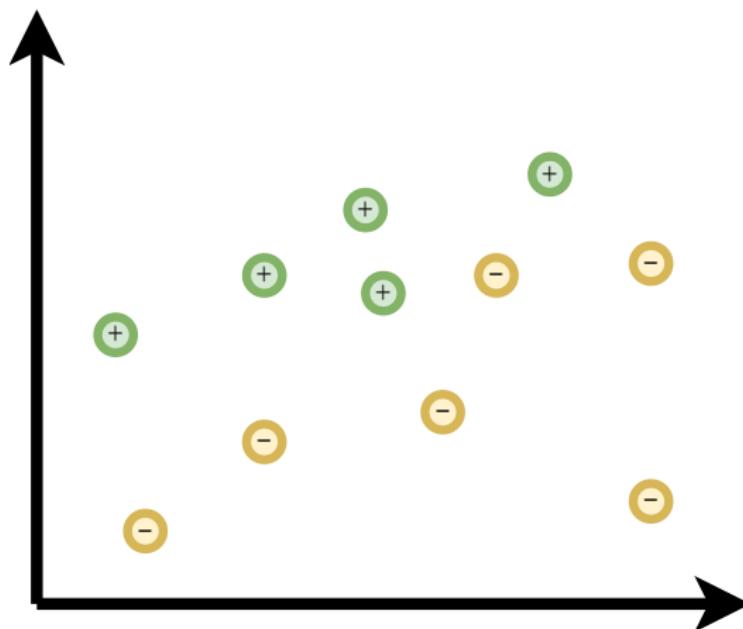


Figura: Exemplo de kNN.

k-Nearest Neighbors

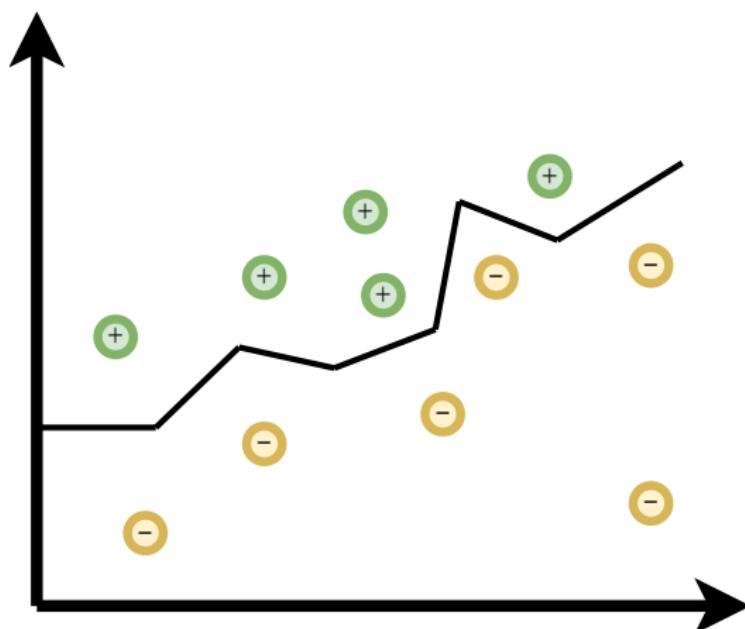


Figura: Exemplo de kNN.

k-Nearest Neighbors

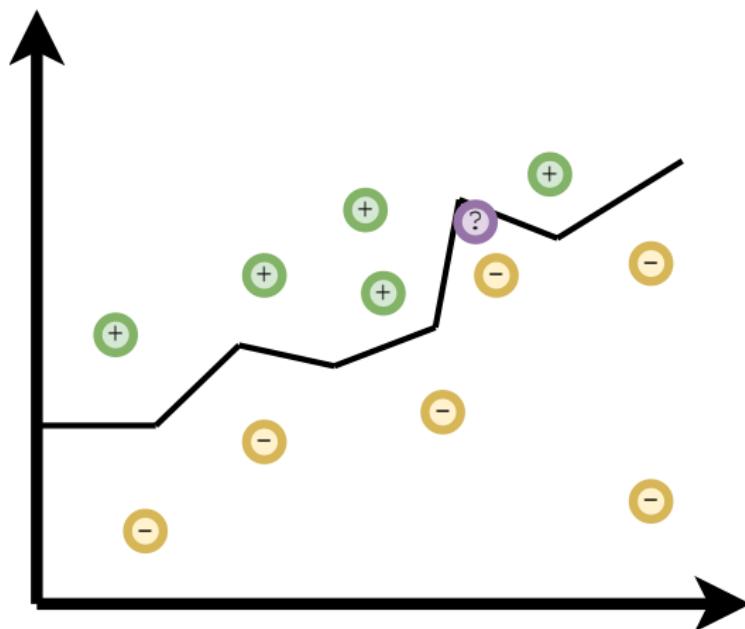


Figura: Exemplo de kNN.

k-Nearest Neighbors

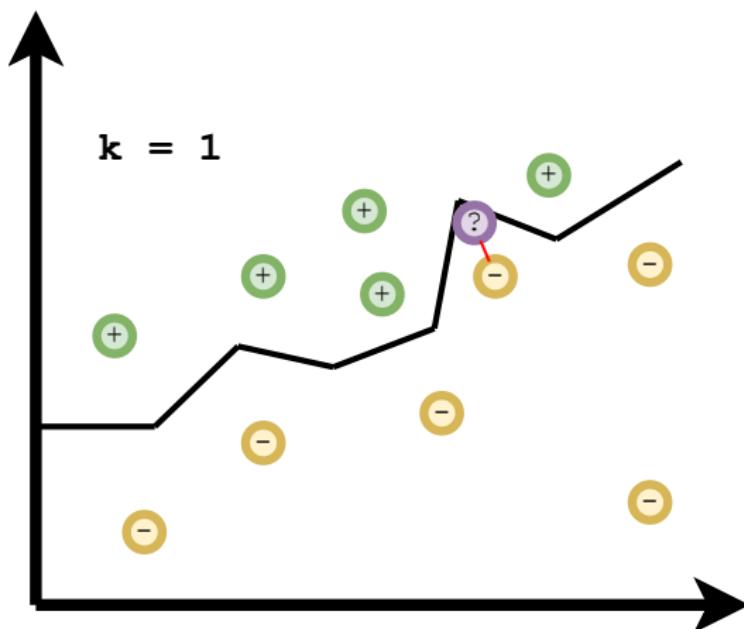


Figura: Exemplo de kNN.

k-Nearest Neighbors

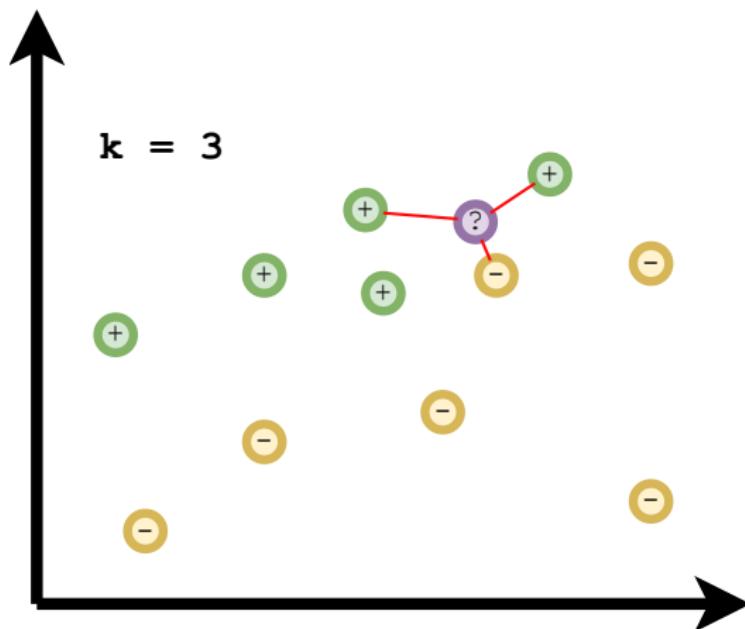


Figura: Exemplo de kNN.

k-Nearest Neighbors

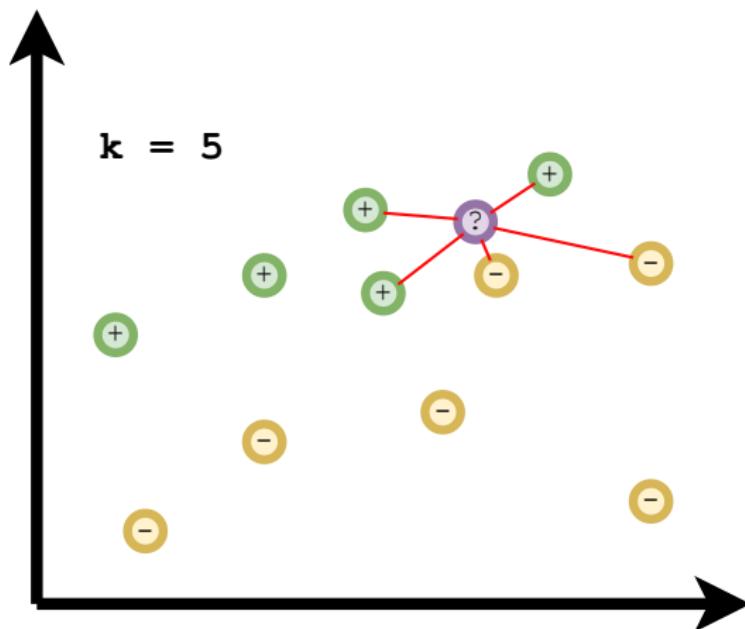


Figura: Exemplo de kNN.

k-Nearest Neighbors

- Vantagens

- Não possui etapa de treinamento
- Processo de teste extremamente simples (distância euclidiana)
- Poucos hiperparâmetros a serem otimizados

- Desvantagens

- Todas as amostras são salvas no modelo
- Para o teste, uma busca deve ser feita com complexidade $\Theta(N)$
- Problemas com generalização e outliers

Decision Trees

- Naturalmente hierárquico
- Computacionalmente barato
- Algoritmo recursivo simples de se implementar

Exemplo de Decision Tree

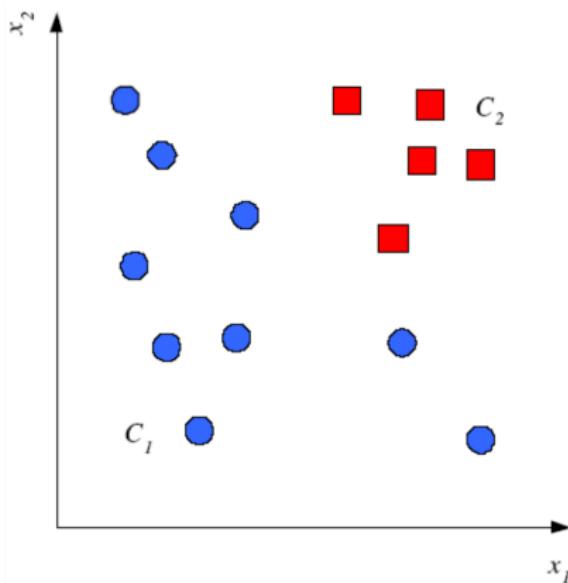


Figura: Exemplo de Decision Tree.

Exemplo de Decision Tree

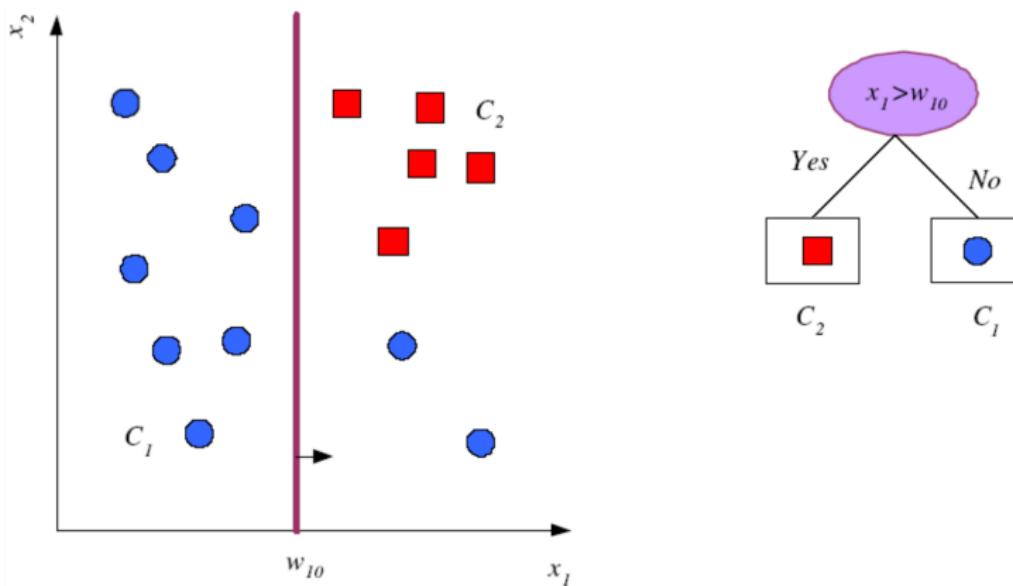


Figura: Exemplo de Decision Tree.

Exemplo de Decision Tree

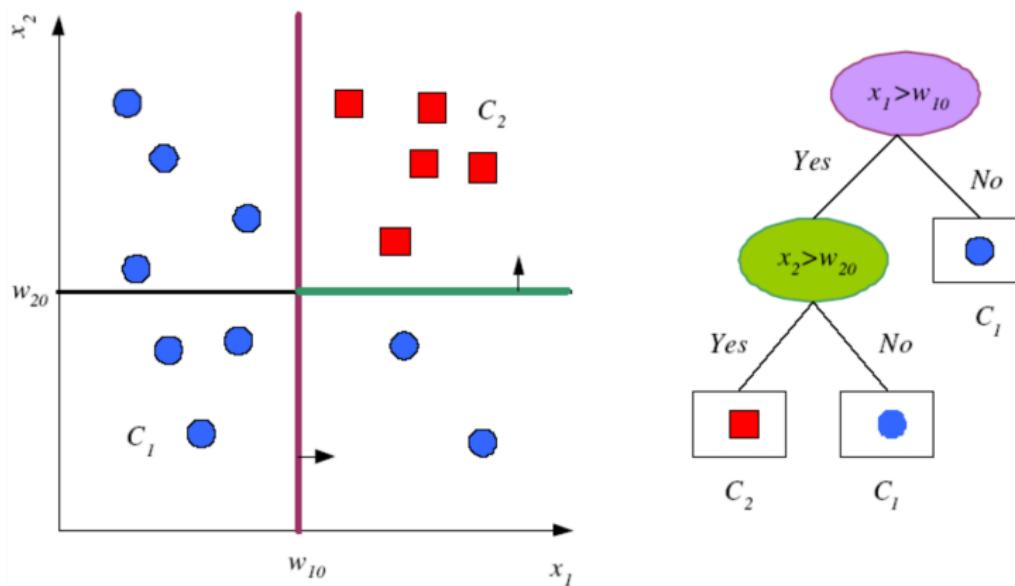


Figura: Exemplo de Decision Tree.

Exemplo de Decision Tree

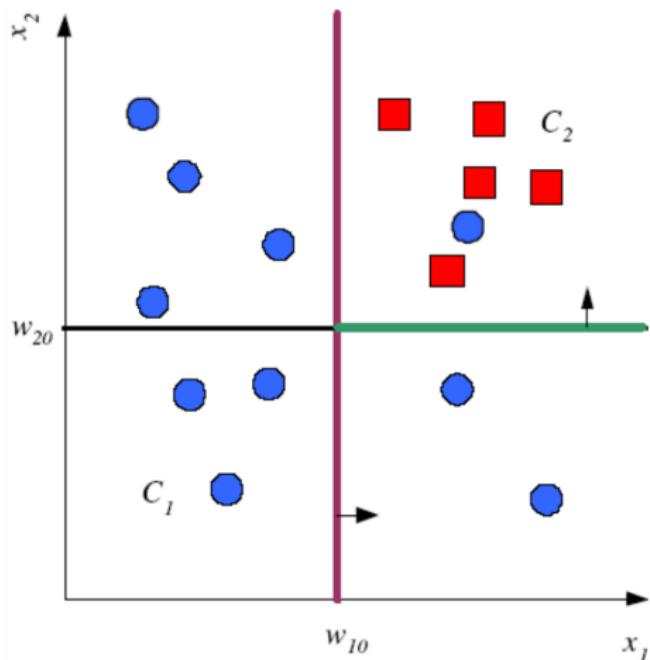


Figura: Exemplo de Decision Tree.

Exemplo de Decision Tree

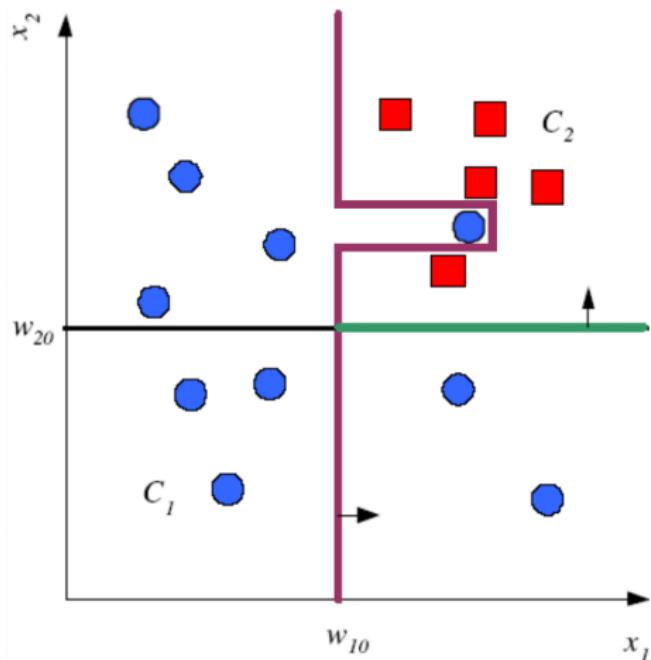


Figura: Exemplo de Decision Tree.

Decision Trees

Complexidade em Decision Trees

É possível enviesar o modelo controlando a altura máxima da árvore.

Controlando a Complexidade das DTs

Creativision

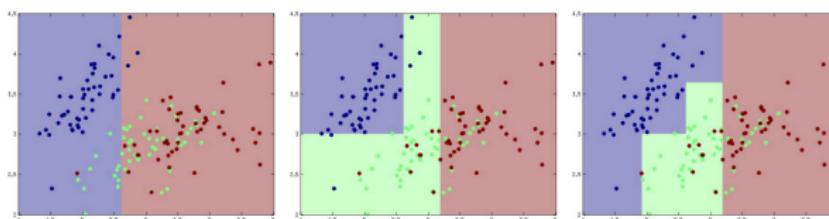
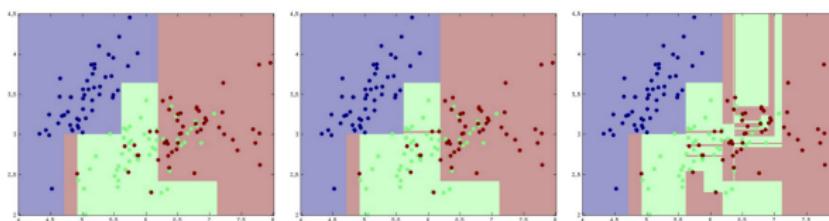
(a) $d_{max} = 1$ (b) $d_{max} = 2$ (c) $d_{max} = 3$ (d) $d_{max} = 4$ (e) $d_{max} = 5$ (f) $d_{max} = \infty$

Figura: Complexidade das DTs.

Decision Trees

- Early Stop é normalmente utilizado para evitar overfit
 - Caso o modelo não seja impedido de crescer, ruído nos dados facilmente pode levar a overfit
 - Um modelo é considerado bom e para de crescer quando os nós atingem um certo grau de “pureza”

Decision Trees

Pureza de Nós em uma Decision Tree

Como medir pureza nos nós de uma Decision Tree?

Decision Trees

Pureza de Nós em uma Decision Tree

Como medir pureza nos nós de uma Decision Tree?

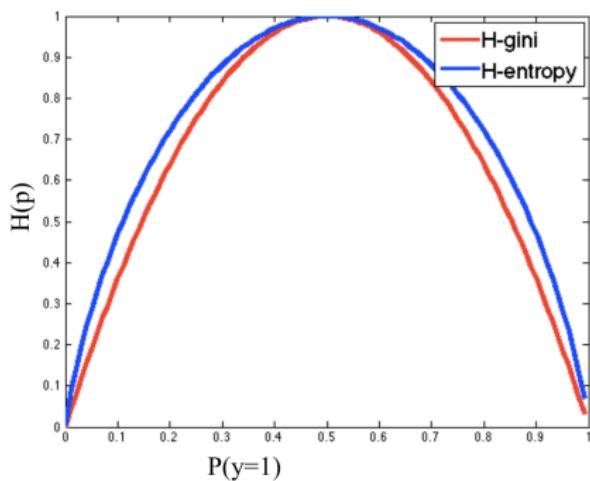


Figura: Pureza nos nós de uma Decision Tree.

Decision Trees

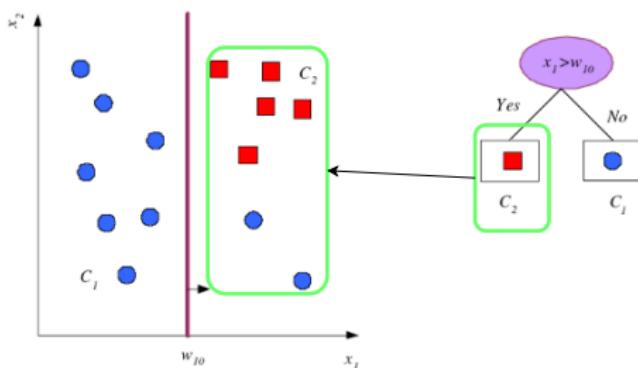


Figura: Pureza nos nós de uma Decision Tree.

- Classe vermelha: $\frac{5}{7} \approx 0.71$
- Classe azul: $\frac{2}{7} \approx 0.29$

Decision Trees

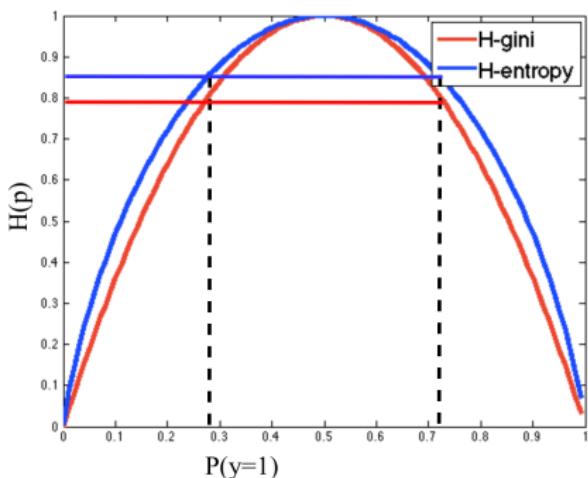


Figura: Pureza nos nós de uma Decision Tree.

- Gini: $G(p) = 1 - \sum_{i=1}^J p_i^2$
- Entropia: $H(p) = -\sum_{i=1}^J p_i \log_2 p_i$

Support Vector Machines

- Boa capacidade de generalização (maximização da margem)
- Robusto para dados com muitas dimensões
- Forte background matemático
- Possibilidade do uso de kernels
- Modelo bastante compacto após treinamento

Support Vector Machines

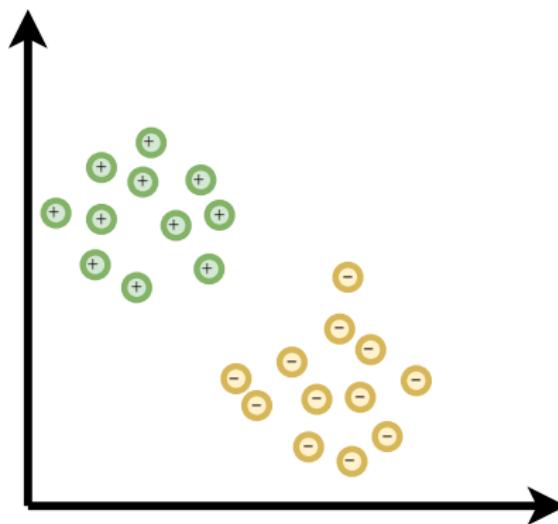


Figura: Classificação usando Support Vector Machines.

Support Vector Machines

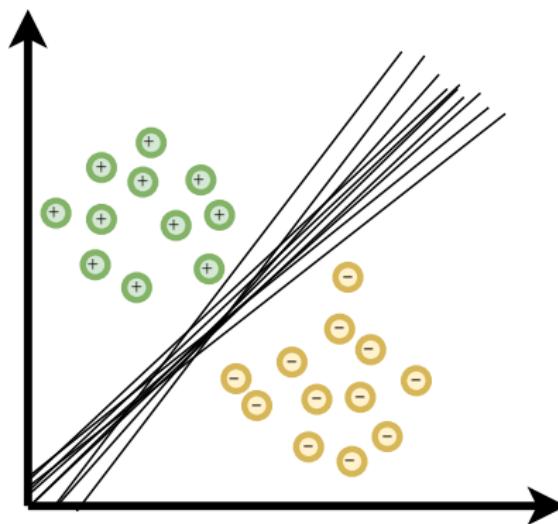


Figura: Classificação usando Support Vector Machines.

Support Vector Machines

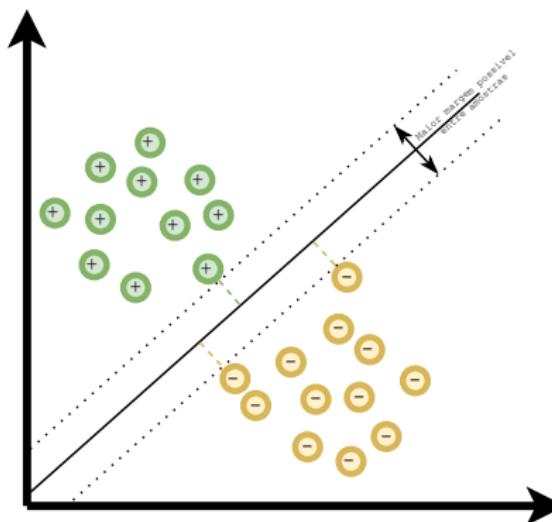


Figura: Classificação usando Support Vector Machines.

Support Vector Machines

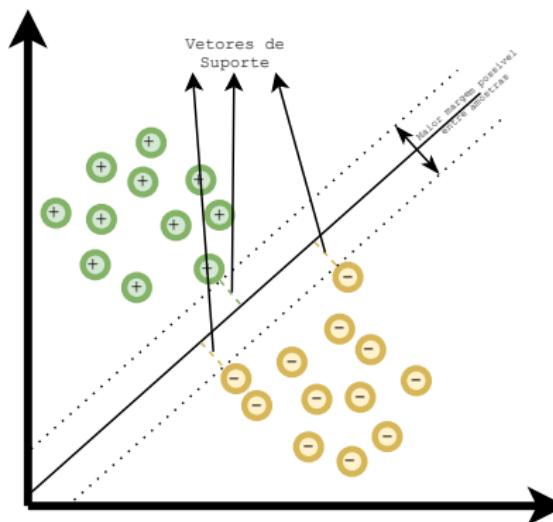


Figura: Classificação usando Support Vector Machines.

Support Vector Machines

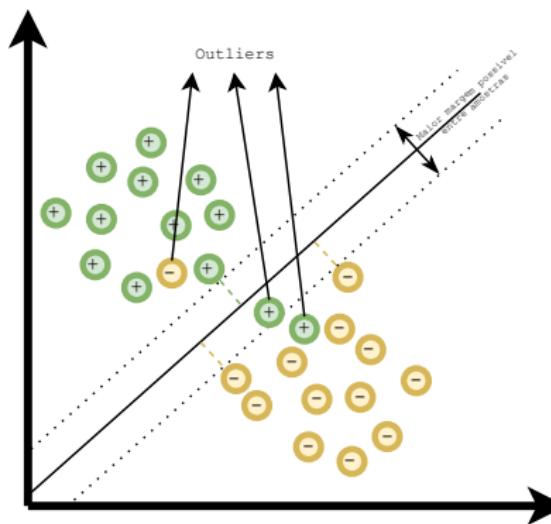


Figura: Classificação usando Support Vector Machines.

Kernels em Support Vector Machines

- Muitas vezes os dados de treinamento não são linearmente separáveis
- SVMs pertencem a um conjunto de algoritmos que podem contar com o uso de kernels
- Várias funções diferentes podem servir como kernels
 - Funções Polinomiais
 - Radial Basis Function (RBF)
 - Tangente Hiperbólica (tanh)

Kernels

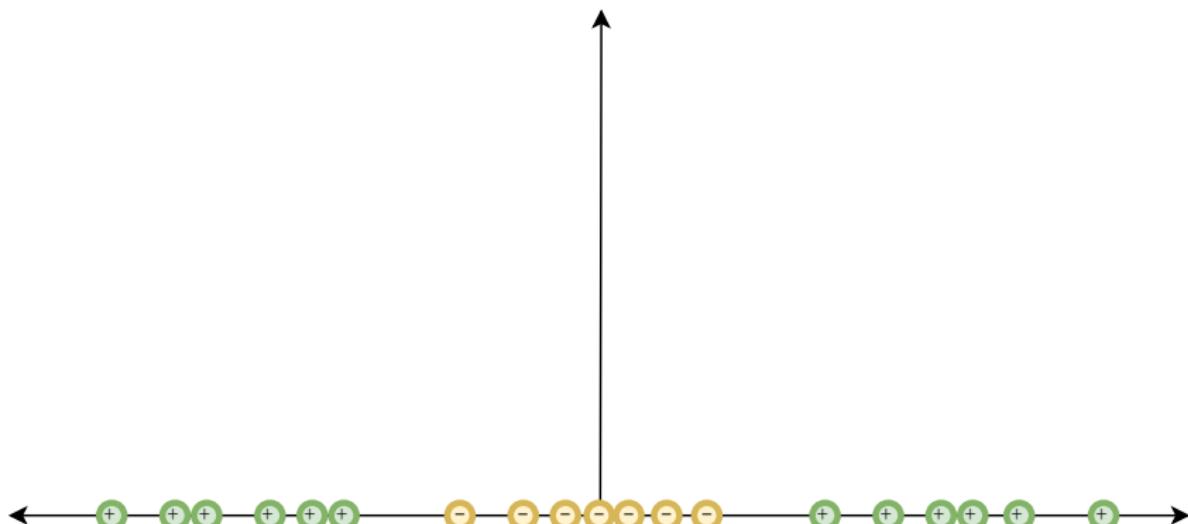


Figura: Kernels em Support Vector Machines.

Kernels

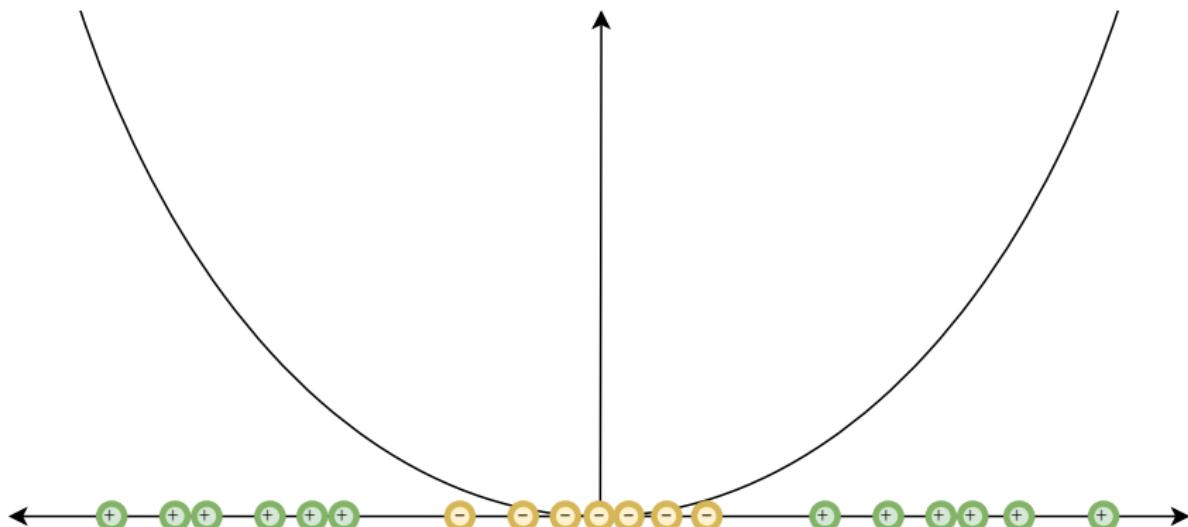


Figura: Kernels em Support Vector Machines.

Kernels

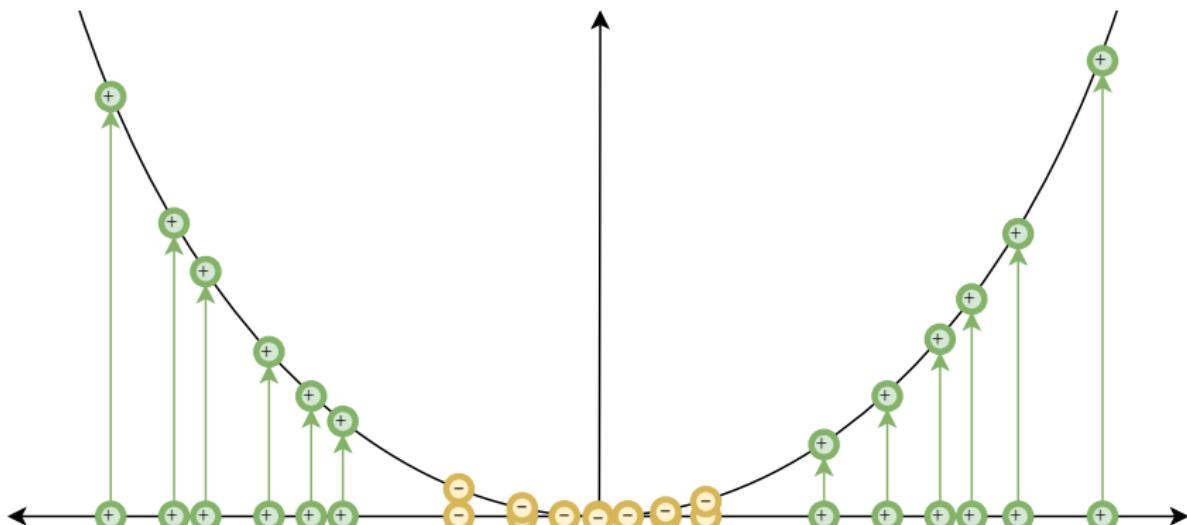


Figura: Kernels em Support Vector Machines.

Kernels

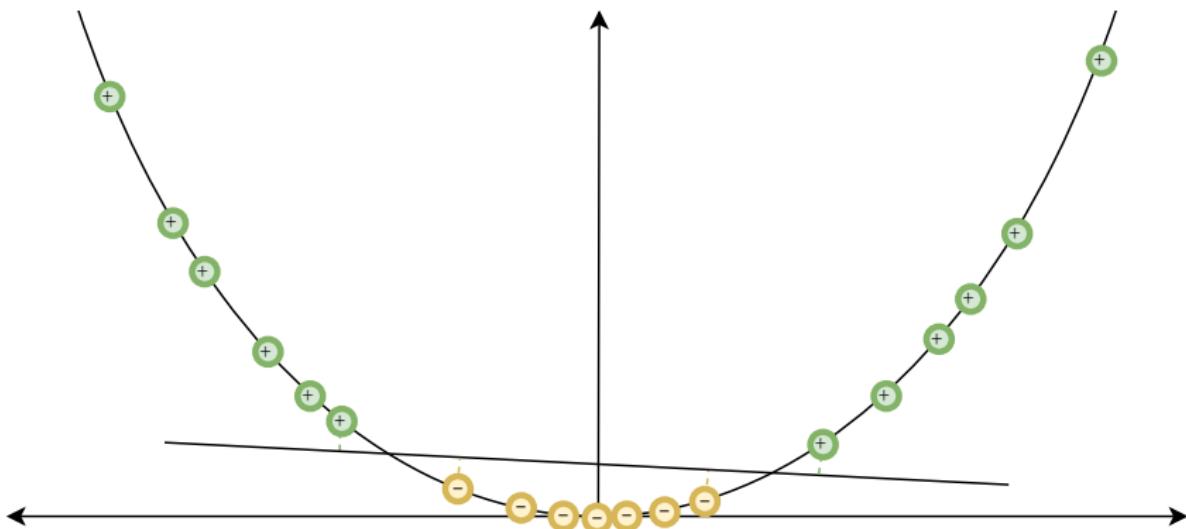


Figura: Kernels em Support Vector Machines.

SVM Kernels

Demo (Extra) - Kernels do SVM

svm_kernels.ipynb

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Avaliação de Modelos

Avaliação de Modelos

Como garantir que meu modelo está com um bom trade-off entre viés e variância?

Avaliação de Modelos

Avaliação de Modelos

Como garantir que meu modelo está com um bom trade-off entre viés e variância?

- É preciso quantificar a quantidade de erros de predição que o modelo de inferência gerou
- Métricas de erro

Avaliação de Modelos

- Métricas para problemas de Clusterização
 - Homogeneidade
 - Informação Mútua
- Métricas para problemas de Regressão
 - R^2
 - Mean Squared Error (MSE)
- Métricas para problemas de Classificação
 - Acurácia
 - F1 (Dice) score
 - Curva ROC/AUC (problemas binários)

Protocolos de Avaliação

- Conjuntos de Dados
 - Conjunto de Treino
 - Usado para treinar os modelos
 - Conjunto de Teste
 - Usado para avaliar os modelos
 - Conjunto de Validação
 - Usado para tunar os parâmetros dos modelos aos dados

Protocolos de Avaliação



Figura: Divisão entre Treino, Validação e Teste.

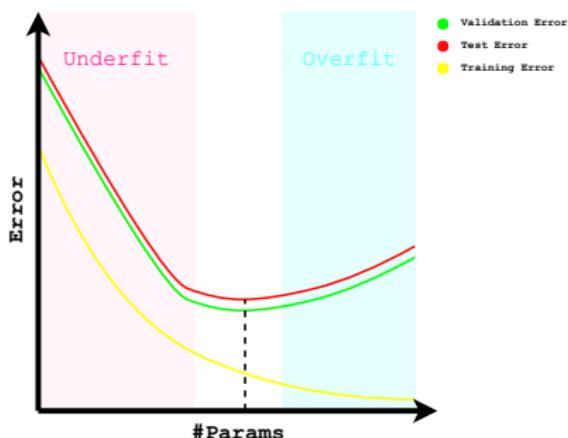


Figura: Erro vs. #Params.

Organização de Dados no sklearn

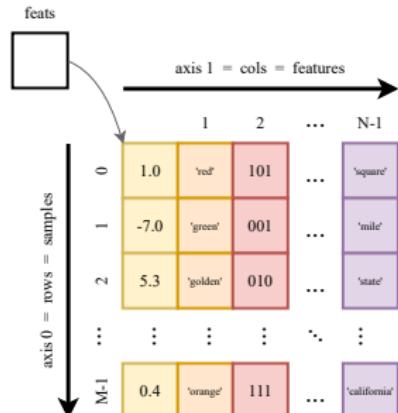


Figura: Organização da matriz de características.

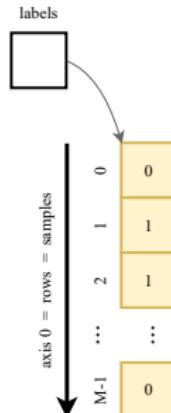


Figura: Organização do vetor de labels (se preciso).

Protocolo de Treino/Teste Shallow

Demo - Pipeline de Treino/Teste de Algoritmos Shallow

shallow_pipeline_toy_datasets.ipynb

Protocolo de Treino/Teste Shallow

Prática - Treinando um Reconhecedor de Dígitos

shallow_pipeline_mnist.ipynb

Aprendizado Visual Clássico

Limitações do Aprendizado Shallow

Desenhar características relevantes para cada tarefa em particular é um problema muito difícil, senão impossível. Para a maior parte das tarefas de reconhecimento de imagens no mundo real, as features existentes de cor, textura e forma não são discriminativas ou representativas o bastante.

¹ <https://hal.archives-ouvertes.fr/file/index/docid/625533/filename/1569408599.pdf>

² https://link.springer.com/chapter/10.1007/978-3-642-25382-9_2

³ <https://dl.acm.org/doi/abs/10.1145/1290082.1290111>

Aprendizado Visual Clássico

Limitações do Aprendizado Shallow

Desenhar características relevantes para cada tarefa em particular é um problema muito difícil, senão impossível. Para a maior parte das tarefas de reconhecimento de imagens no mundo real, as features existentes de cor, textura e forma não são discriminativas ou representativas o bastante.

Como projetar features mais adaptadas às tarefas individuais?

Uma forma de mitigar esse problema é desenhar extractores de características que se adequem automaticamente aos dados da tarefa. Durante as décadas de 2000 e o começo da de 2010, abordagens de Bag of Visual Words (BoVW)¹²³ se propuseram a resolver esse problema.

¹ <https://hal.archives-ouvertes.fr/file/index/docid/625533/filename/1569408599.pdf>

² https://link.springer.com/chapter/10.1007/978-3-642-25382-9_2

³ <https://dl.acm.org/doi/abs/10.1145/1290082.1290111>

Agenda

1 Introdução

- O Pipeline de um Algoritmo de Aprendizado Shallow

2 Extração de Features

3 Shallow Learning

- Algoritmos de Aprendizado de Máquina
- Avaliação de Modelos

4 Bag of Visual Words

Aprendizado Visual ao Longo das Décadas

Creativision

Input



Output



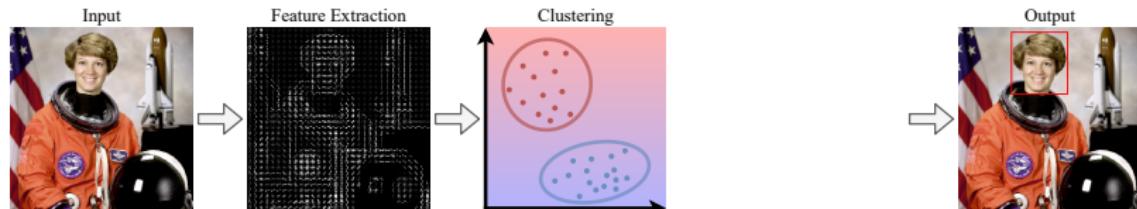
Aprendizado Visual ao Longo das Décadas

Creativision



Aprendizado Visual ao Longo das Décadas

Creativision

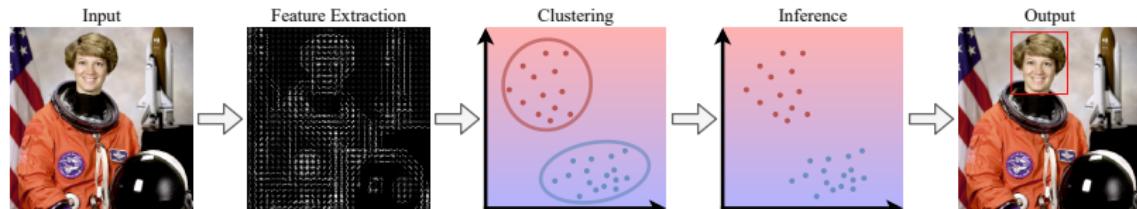


Bag of Visual Words (BoVW)

Ao invés de usar diretamente as características locais ou globais das imagens, abordagens de BoVW se utilizam de algoritmos de clustering para aprender features de nível semântico um pouco mais próximo dos dados em que estão operando. O intuito do clustering é achar grupos de amostras mais e menos correlacionadas no espaço de features.

Aprendizado Visual ao Longo das Décadas

Creativision



Visual Recognition with Bags of Visual Words (BoVW)

Normalmente algum tipo de histograma sobre a atribuição de clusters (i.e. soft ou hard) ou uma distância tensorial para o centroide do cluster é a informação passada para o algoritmo de aprendizado (i.e. SVM, RF, NN, etc) ao invés de passar diretamente as features cruas.

Aprendizado Visual ao Longo das Décadas

Limitações do BoVW

Mesmo os features adaptáveis para cada task obtidos pelo processo de clustering ainda dependem da capacidade de representatividade dos extractores de características escolhidos, sendo esse o maior gargalo de desempenho nas abordagens baseadas em BoVW. Hoje em dia na era do Deep Learning é consideravelmente difícil de se encontrar algum tipo de abordagem baseada em BoVW.

Aprendizado Visual ao Longo das Décadas

Limitações do BoVW

Mesmo os features adaptáveis para cada task obtidos pelo processo de clustering ainda dependem da capacidade de representatividade dos extratores de características escolhidos, sendo esse o maior gargalo de desempenho nas abordagens baseadas em BoVW. Hoje em dia na era do Deep Learning é consideravelmente difícil de se encontrar algum tipo de abordagem baseada em BoVW.

Abandonando features desenhados manualmente

A solução para essa limitação foi encontrada no Deep Learning, começando em 2012, apesar de Redes Neurais já contarem com mais de 50 anos de estudo antes dessa data. No Deep Learning **as features são aprendidas em conjunto com a otimização do algoritmo de inferência**.

└ References

- [1] NVIDIA.
Deep learning - nvidia developer, 2018.
- [2] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio.
Deep learning, volume 1.
MIT press Cambridge, 2016.

