

Aprendizagem automática

Introdução

Luís Rato, Universidade de Évora, 2022/23
Com base em AA 21/22 Teresa Gonçalves

Sumário

Introdução

Conceitos básicos

Aprendizagem profunda

Conjuntos, Desafios e Frameworks

Bibliografia

Introdução

O que é a Aprendizagem Automática ?

Onde se situa a aprendizagem automática nos **sistemas de raciocínio**?

O raciocínio/inferência divide-se em duas grandes abordagens (Aristóteles 300 AC):

#1 - Raciocínio dedutivo - Pressupostos, Regras, Conclusões - **Programação em Lógica**

#2 - Raciocínio indutivo* - Observações, Conclusões - **Aprendizagem**

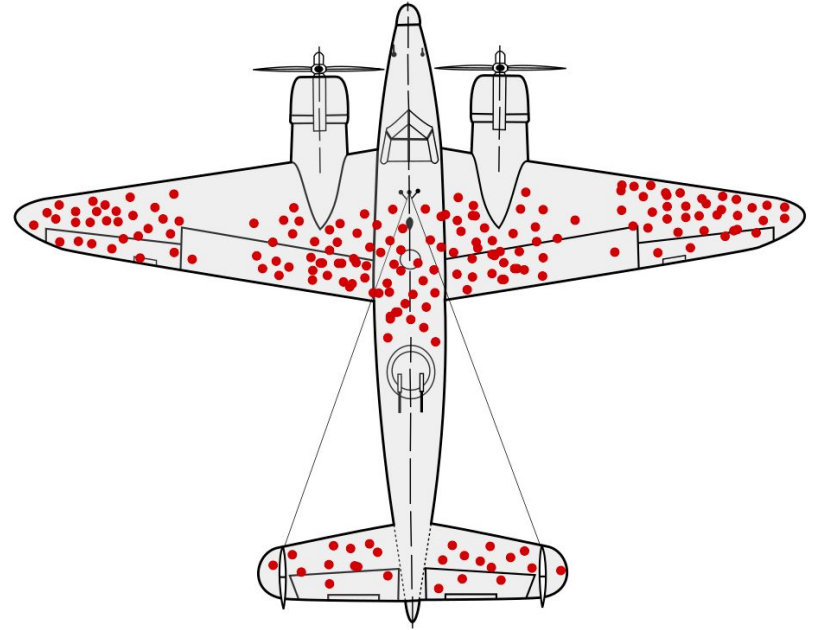
*pode ainda referir-se o raciocínio **abduativo** podendo ser visto como um raciocínio indutivo sobre a conclusão mais plausível.

O que é a Aprendizagem Automática ?

Falácia: é um erro de raciocínio.

Em AA temos de lidar com os erros dos modelos preditivos.

É muito importante não introduzir mais erros



Survival Bias

O que é a Aprendizagem Automática ?

Estudo sistemático de **algoritmos** e sistemas que **otimizam** o seu comportamento com a **experiência** (ou seja em função dos **dados**)

*Do inglês, *Machine Learning*

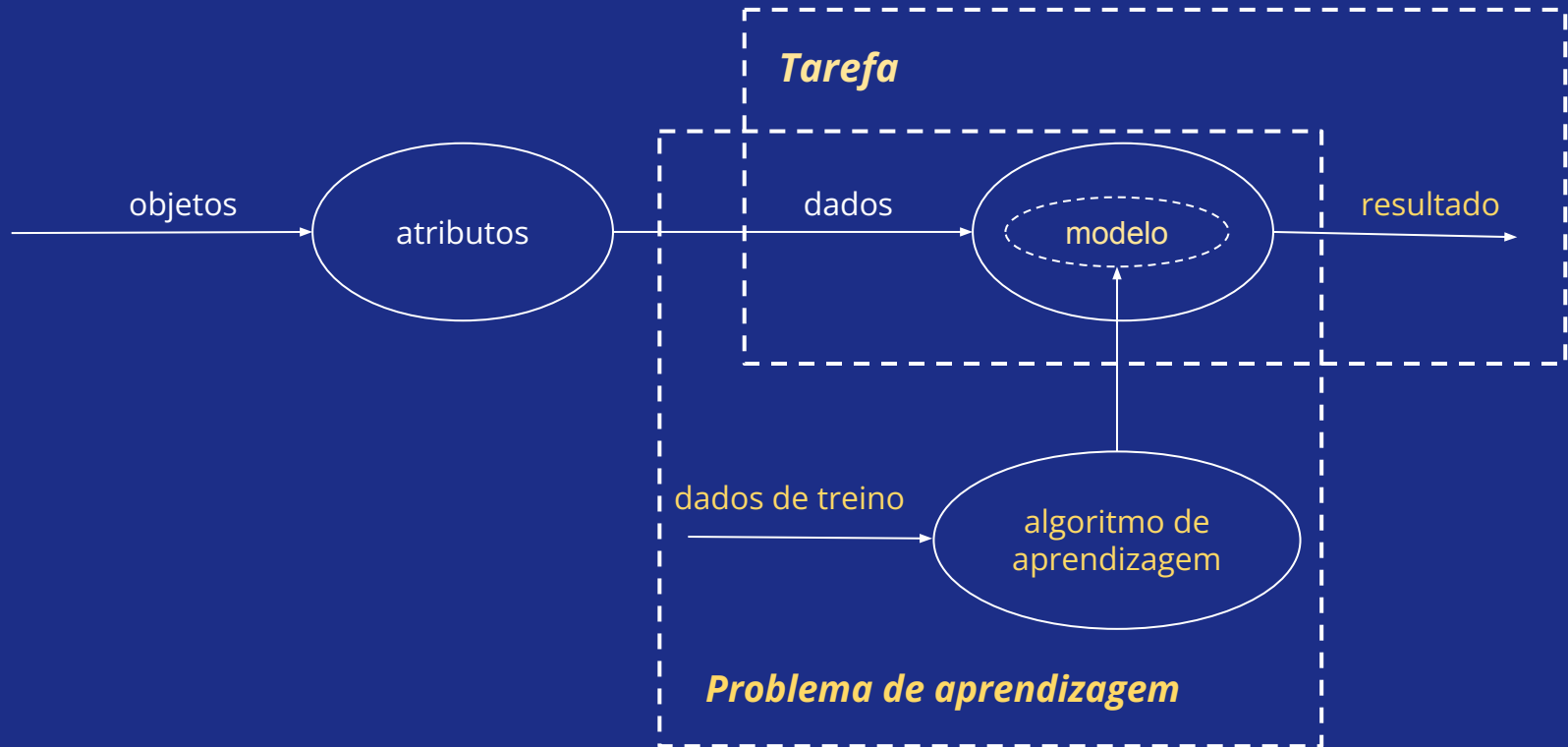
Aplicações de Aprendizagem Automática

- Assistente pessoal virtual
 - Siri, Alexa, Google Now
- Serviços financeiros
 - Detecção de fraudes
- Cuidados de Saúde
 - Análise da condição do paciente a partir de sensores wearable
 - Reconhecimento de tecidos cancerosos
 - Diagnóstico de doenças raras por reconhecimento facial
 - Identificação de sintomas do paciente usando chatbots
- Marketing e vendas
 - Sugestão de produtos
 - Amazon, Spotify, Netflix
- Trabalho diário
 - Marcação de mensagens Spam
 - Filtro de Malware
- Redes sociais
 - Sugestão de amigos
- Serviços de texto
 - Análise de textos
 - Processamento de língua natural
 - Tradução automática
 - Google Translate
- Visão computacional
 - Reconhecimento de imagens
 - Google Photos
 - <https://clarifai.com/models>

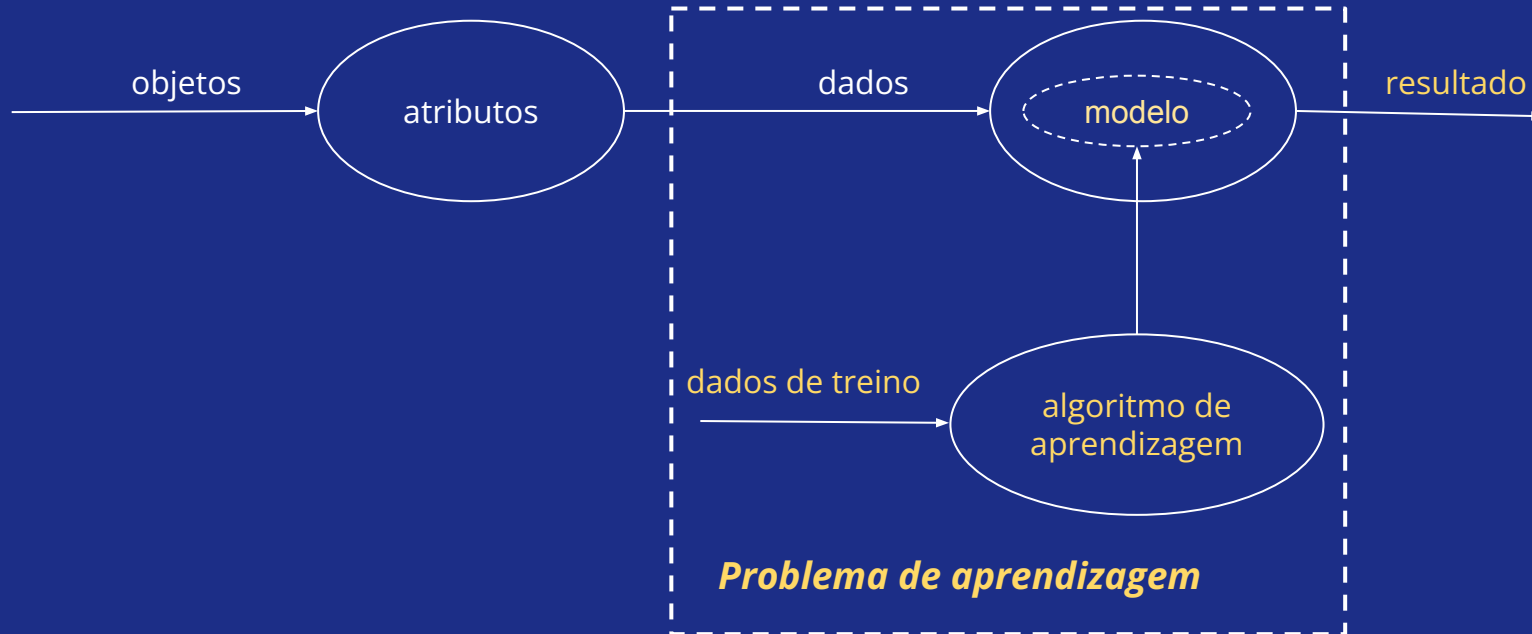
Visão geral de um sistema de AA



Visão geral de um sistema de AA



Visão geral de um sistema de AA



Conceitos básicos

Conceitos a reter

- Generalização
 - Capacidade do sistema generalizar **para além** da experiência (exemplos vistos)
 - Provavelmente o conceito mais **fundamental** em AA
- Sobre-ajustamento (*overfitting*)
 - Quando o sistema memoriza a experiência
 - **Não existe generalização** ou é pobre
- Avaliação do sistema
 - Para ter uma ideia de **quão bem o modelo irá funcionar** com novos dados
- No free lunch theorem
 - **Nenhum** algoritmo de aprendizagem **é melhor** que outro quando avaliados sobre **todos os problemas** possíveis

Terminologia AA

- **Objetos**
 - Unidades de entrada de um sistema AA
 - Também conhecido como **instâncias (ou elementos, pontos, etc...)**
- **Atributos**
 - Descritores dos objetos
 - Também conhecidos como **características**
- **Saída**
 - O que o modelo associa ao objeto de entrada
 - Pode ser um grupo, classe, número inteiro ou número real, ...

Ingredientes principais de AA

- Tarefas
 - São **representações** abstratas dos problemas que queremos resolver
- Modelos
 - São o **resultado** de um algoritmo de AA aplicado aos dados de treino
- Atributos
 - Definem as características dos objetos, definem uma **linguagem** para descrever os objetos relevantes do domínio

*AA interessa-se em usar
os **atributos** certos
para construir os **modelos** certos
que realizam as **tarefas** certas*

Tarefas: os problemas que podem ser resolvidos

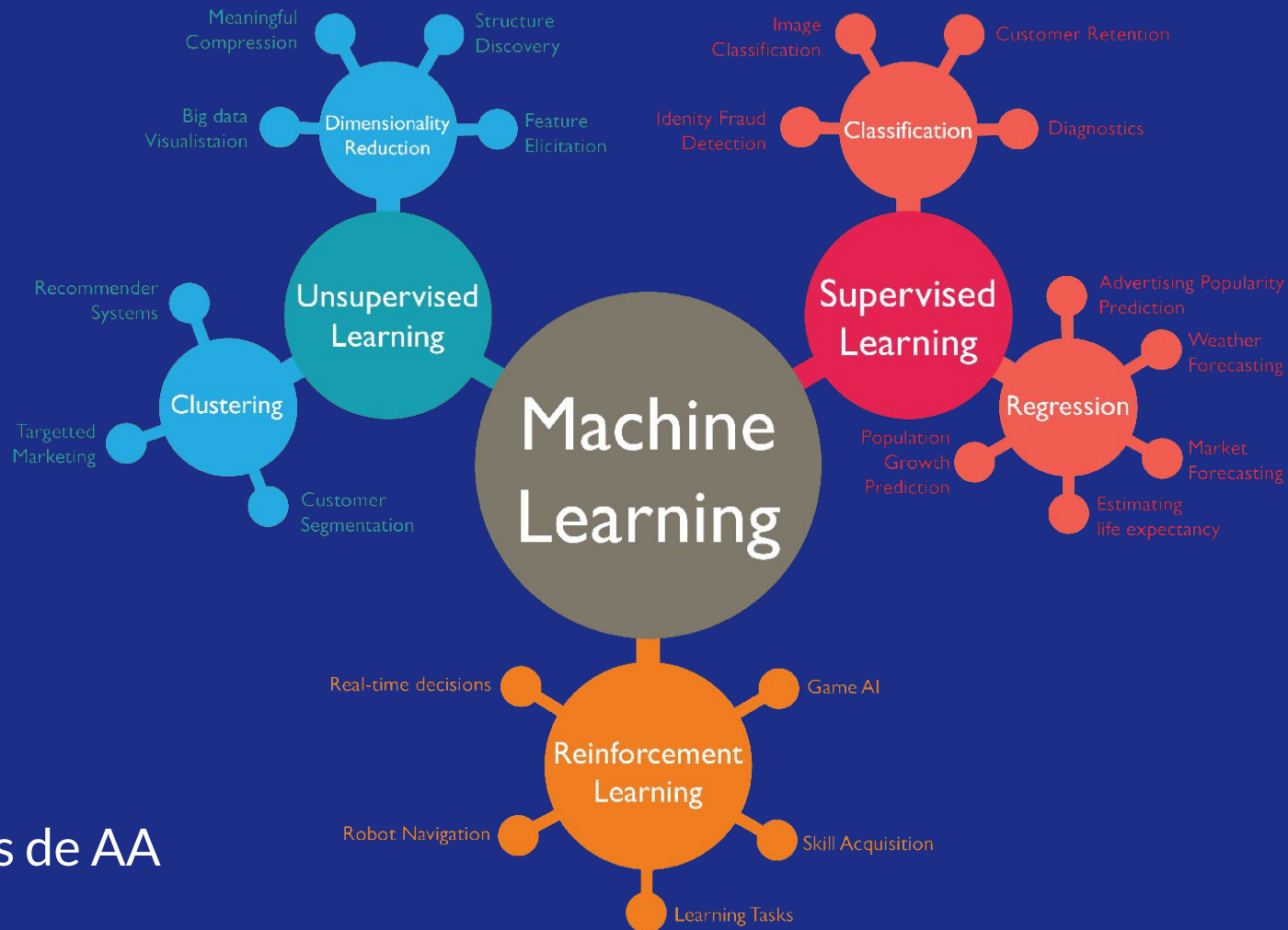
- Classificação
 - Prevê um conjunto de classes discretas
- Regressão
 - Prevê um número real
- Agrupamento
 - Agrupa dados (sem informação priori sobre os grupos)
- Descoberta de associações
 - Encontra associações entre itens que co-ocorrem frequentemente

**Aprendizagem
supervisionada**

**Aprendizagem
não
supervisionada**

Exemplos

- Classificação
 - Distinguir entre mensagens de spam e não spam (sim/não)
- Regressão
 - Avaliar a urgência de uma mensagem de mail (escala 0-10; 0=ignorar, 10=ação imediata)
- Agrupamento
 - Colocar mensagens semelhantes no mesmo grupo e diferentes mensagens em grupos diferentes
- Descoberta de associações
 - Encontrar os itens em cestos de compras que ocorrem com frequência



Tarefas de AA

Modelos: o resultado da AA

- Escolha de uma ampla variedade de modelos para escolher
- Algoritmos desenvolvidos por
 - matemáticos
 - engenheiros
 - psicólogos
 - Cientistas da computação
 - outros
- Grupos de modelos
 - Geométricos
 - Probabilísticos
 - Lógicos

Modelos Geométricos

- Ideia
 - O espaço de instâncias tem uma estrutura geométrica
 - Para atributos numéricos, pode ser usado o sistema cartesiano de coordenadas. Cada atributo é uma coordenada
- Modelo construído diretamente no espaço de instâncias
 - Utiliza conceitos geométricos como linhas, planos, e distâncias
- Vantagens
 - Fácil visualização (para 2 ou 3 dimensões)

Algoritmos de AA



Atributos: os “trabalhadores” da AA

- Medições realizadas sobre uma instância
 - Pode ser visto como uma função que **mapeia** o **espaço de instâncias** no **domínio dos atributos** (o conjunto de valores possíveis dos atributos)
- Também conhecidos como **características**, ou **Variáveis preditivas**, ou **Variáveis explicativas**

Variáveis independentes

*Os atributos determinam muito do **sucesso** de uma aplicação de AA!*

Tipos de atributos

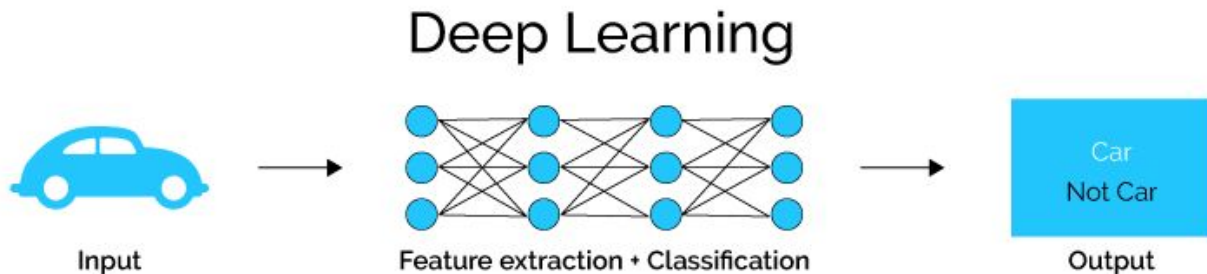
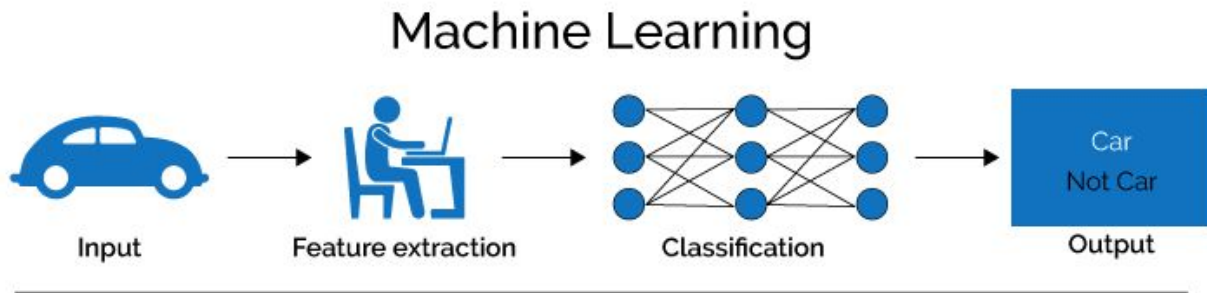
- Simples
 - As instâncias são representadas por um vetor de valores
 - Domínio
 - Real, conjunto de inteiros
 - Booleano, conjunto finito arbitrário
- Estruturados
 - Sequências (ex: strings, sequência de palavras)
 - Árvores (ex: estrutura HTML)
 - Grafos (ex: redes sociais)

Engenharia de atributos

- Objetivo
 - Encontrar o **melhor conjunto** de atributos para a tarefa
- Operações
 - Transformação
 - Alterar um atributo original através da remoção, alteração ou adição de informação
 - Construção
 - Construir novos atributos a partir de um conjunto de atributos originais
 - Seleção
 - Selecionar um sub-conjunto apropriado antes da aprendizagem
 - Aumentar a rapidez da aprendizagem (considerando menos atributos)
 - Ajudar na proteção contra o sobre-ajustamento

Aprendizagem Profunda

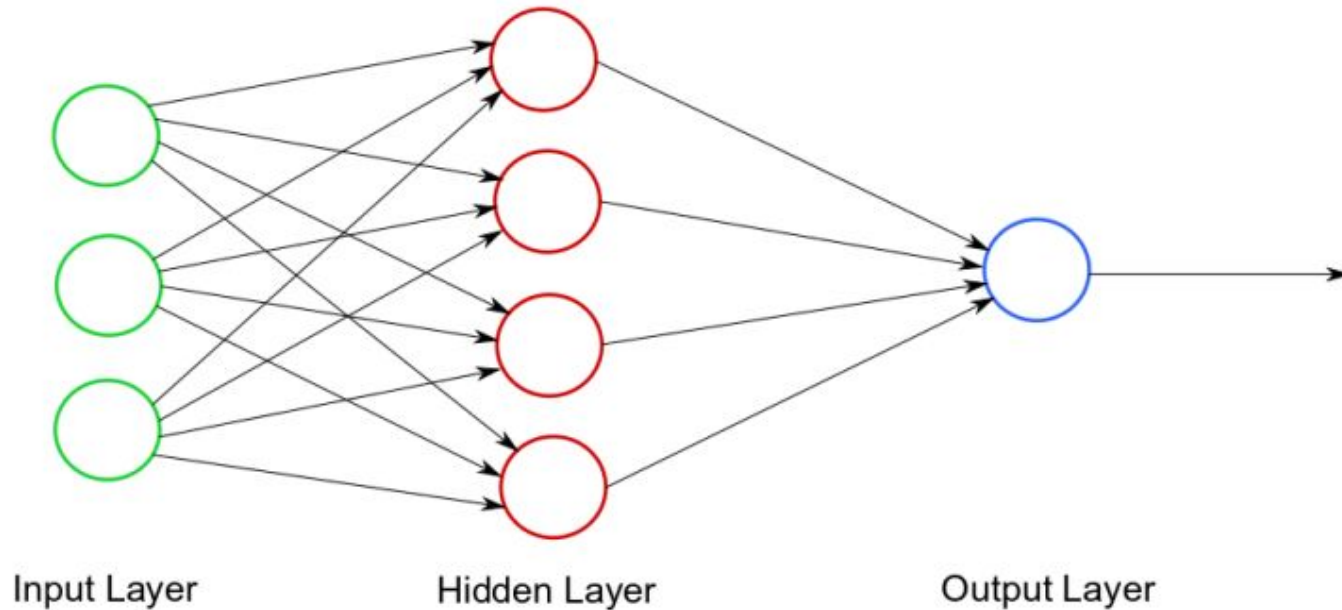
Aprendizagem Automática vs. Aprendizagem profunda



Aprendizagem profunda

- Algoritmo
 - Rede neuronal artificial (RNA) com várias camadas escondidas
 - A arquitetura da RNA depende da aplicação
- Quando usar?
 - Quando existe uma ENORME quantidade de exemplos

Rede Neuronal Artificial



Tipos de redes / arquiteturas

- FFNN - Redes feedforward - Uma camada liga-se apenas à camada seguinte
- RNN - Redes recorrentes - Permite ligações em ciclo entre neurónios
- CNN - Convolutional Neural Network
- LSTM - Long Short-term Memory Network
- GAN - Generative Adversarial Network

Aplicações

Deep Dream (Google)



AtomNet (Atomwise)

AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery

Izhar Wallach
Atomwise, Inc.
izhar@atomwise.com

Michael Dzamba
Atomwise, Inc.
misk@atomwise.com

Abraham Heifets
Atomwise, Inc.
abe@atomwise.com

Abstract

Deep convolutional neural networks comprise a subclass of deep neural networks (DNN) with a constrained architecture that leverages the spatial and temporal structure of the domain they model. Convolutional networks achieve the best predictive performance in areas such as speech and image recognition by hierarchically composing simple local features into complex models. Although DNNs have been used in drug discovery for QSAR and ligand-based bioactivity predictions, none of these models have benefited from this powerful convolutional architecture. This paper introduces AtomNet, the first structure-based, deep convolutional neural network designed to predict the bioactivity of small molecules for drug discovery applications. We demonstrate how to apply the convolutional concepts of feature locality and hierarchical composition to the modeling of bioactivity and chemical interactions. In further contrast to existing DNN techniques, we show that AtomNet's application of local convolutional filters to structural target information successfully predicts new active molecules for targets with no previously known modulators. Finally, we show that AtomNet outperforms previous docking approaches on a diverse set of benchmarks by a large margin, achieving an AUC greater than 0.9 on 57.8% of the targets in the DUDC benchmark.

AlphaGo (Google)



Aplicações

GAN - “the coolest idea in Machine Learning in the last twenty years!”

Envelhecimento facial (Facebook)



Exemplo de progressão da capacidade das GANs



*Retirado de “The Malicious Use of Artificial Intelligence:
Forecasting, Prevention, and Mitigation”, 2018*

Conjuntos de dados,
Frameworks

Desafios e

Conjuntos de dados e desafios

- Conjuntos gerais
 - UCI: <https://archive.ics.uci.edu/ml/datasets.php>
- Ciência dos dados
 - Kaggle: <https://www.kaggle.com/>
- Imagens
 - ImageNet: <http://image-net.org/challenges/LSVRC/>
- Recuperação de informação (texto, imagens, vídeo)
 - CLEF Labs (desde 2003): <http://www.clef-initiative.eu/>
- Recuperação de texto
 - TREC tasks: <https://trec.nist.gov/>
 - NTCIR tasks: <http://research.nii.ac.jp/ntcir/index-en.html>
 - FIRE tasks: <http://fire.irsu.res.in>
 - SemEval: <http://alt.qcri.org/semeval2019/>

Frameworks de AA

- Académico
 - WEKA: <https://www.cs.waikato.ac.nz/ml/weka/>
 - Vários interfaces: CLI, Explorer, Experimenter, KnowledgeFlow
- Mais orientado ao negócio
 - RapidMiner: <https://rapidminer.com/>
 - Duas versões: livre (versão community) e paga
- Python
 - Scikit-Learn: <https://scikit-learn.org>
 - Construído sobre NumPy, SciPy, e matplotlib
- Outros
 - Desenvolvidos com propósitos específicos (texto, imagens, etc...)

Frameworks de Aprendizagem Profunda

- TensorFlow
 - <https://www.tensorflow.org>
 - Uma plataforma open-source end-to-end para Aprendizagem Automática
- CNTK (Microsoft)
 - <https://docs.microsoft.com/en-us/cognitive-toolkit/>
 - toolkit open-source para aprendizagem profunda comercial distribuida
- KERAS
 - <https://keras.io/>
 - API de alto nível para redes neuronais



Bibliografia

Livros

- *Machine Learning: The Art and Science of Algorithms that make Sense of Data*. Peter Flach. Cambridge University Press, 2012
- *Data Mining: Practical Machine Learning Tools and Techniques*. Ian H. Witten, Eibe Frank, Mark A. Hall and Christopher J. Pal. Morgan Kaufmann, 4th Ed, 2017
- *Machine Learning*. Tom Mitchell. McGraw Hill, 1997
- *Introduction to Machine Learning with Python*. Andreas Muller and Sarah Guido. O'Reilly, 2016

