

Aprendizado de Máquina

Aula 1.2 - Definição e Taxonomia


Adriano Rivolli

rivolli@utfpr.edu.br

Especialização em Inteligência Artificial


Universidade Tecnológica Federal do Paraná (UTFPR)
Câmpus Cornélio Procópio
Departamento de Computação

Conteúdo


- 
- 1 Definição
 - 2 Tarefas de aprendizado
 - 3 Processos relacionados

Definição

Aprendizado (aprender)

- 
- O resultado da atividade de obter conhecimento;
 - O processo de ganhar experiência.


Máquina



Aparelho destinado a produzir, dirigir ou transformar uma forma de energia em outra, ou aproveitar essa mesma energia para a produção de determinado efeito;

Qualquer equipamento empregado com um fim específico e cuja ação mecânica é capaz de substituir o trabalho humano.

Aprendizado de Máquina



*A computer program is said to learn from **experience E** with respect to some class of **tasks T** and performance **measure P** , if its performance at tasks in T , as measured by P , improves with experience E*


— Tom Mitchell

Machine Learning, Tom Mitchell, McGraw Hill, 1997.

Experiência

- Ocorrência de eventos passados
- Objetos que representam um fenômeno
- Exemplos de um caso

Tarefa

- 
- Classificar
 - Estimar um valor
 - Criar um ranking
 - Agrupar
 - Encontrar padrões

Performance

- Acurácia
- Precisão
- Coesão
- Correlação
- Proximidade/Distância
- Tempo

Algoritmo

- Sequência finita e ordenada de instruções bem definidas e não ambíguas
- Programação Imperativa / Orientada a objetos
 - ▶ O algoritmo define como fazer uma tarefa
- Aprendizado de Máquina
 - ▶ O algoritmo define como aprender uma tarefa

Uma segunda definição de AM

*Machine Learning is a branch of artificial intelligence (AI) and computer science which focuses on development of systems that are able to learn and adapt without following **explicit instructions** imitating the way that humans learn, gradually improving its accuracy, by using algorithms and statistical models to analyze and **draw inferences** from **patterns** in data*

— European Union website

Fonte: <https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence>

Raciocínio Dedutivo vs. Indutivo

■ Raciocínio Dedutivo:

- ▶ Parte de premissas gerais para uma conclusão específica
- ▶ Baseado na lógica formal
- ▶ Conclusão necessariamente verdadeira se as premissas forem verdadeiras

■ Raciocínio Indutivo:

- ▶ Parte de exemplos específicos para uma conclusão geral ou probabilidade
- ▶ Baseado em generalizações de padrões observados
- ▶ Conclusão não necessariamente verdadeira, mas provável com base em evidências

Processo Indutivo

■ Observações:

- ▶ Vejo um pássaro e ele tem penas
- ▶ Vejo um novo pássaro e ele também tem penas
- ▶ Vejo um pássaro diferente e ele tem penas
- ▶ ...
- ▶ Todos os pássaros observados até agora têm penas

■ Conclusão Indutiva:

- ▶ Todos os pássaros têm penas.

Quando usar ML?

- Tarefas muito complexas de serem programadas
 - ▶ Tarefas feita por animais/humanos (ex: dirigir, reconhecer a fala)
 - ▶ Tarefas que extrapolam a capacidade humana (ex: predição do tempo)
- Tarefas que precisam se adaptar dinamicamente
 - ▶ Ex: detecção de spam, reconhecimento de texto manuscrito

Tarefas de aprendizado

Tipos de aprendizado

■ Aprendizado Supervisionado

Requer a supervisão para que seja possível aprender a relação entre os dados de entrada e a saída desejada.

■ Aprendizado Não Supervisionado

Não requer supervisão, pois o processo identifica os padrões intrínsecos dos dados.

■ Aprendizado Semi-supervisionado

Combina aprendizado supervisionado com não supervisionado, utiliza poucos dados rotulados.

■ Aprendizado por Reforço

O aprendizado ocorre a partir da interação com o ambiente, utilizando um mecanismo de punição/recompensa.



Aprendizado Supervisionado

■ Classificação

O objetivo é atribuir uma categoria ou rótulo a uma entrada com base em suas características.

■ Regressão

O objetivo é prever um valor numérico com base em uma ou mais variáveis de entrada.

■ Ranking

O objetivo é classificar uma lista de itens de acordo com sua relevância.

Aprendizado Não Supervisionado

■ Agrupamento

O objetivo é organizar os dados em grupos de acordo com algum critério.

■ Redução de dimensionalidade

O objetivo é reduzir o número de variáveis nos dados, mantendo o máximo de informação relevante.

■ Análise de associação

O objetivo é descobrir relações significativas entre variáveis e instâncias.


Processos relacionados

Áreas de atuação

- Análise de dados Interpretação e visualização de dados
- Mineração de dados Aplicação de AM para grandes conjuntos de dados
- Reconhecimento de padrões Aplicação de AM para dados complexos
- Ciência de dados Usa AM para obter insights e comprovar hipóteses
- Inteligência Artificial Campo geral que engloba AM

Como não se perder? Um caminho: <https://i.am.ai/roadmap/>

Processo de AM

- 
- 1 Definição do problema
 - 2 Coleta de dados
 - 3 Pré-processamento
 - 4 Treinamento (etapa que ocorre o aprendizado)
 - 5 Avaliação (pode ser necessário voltar algumas etapas)
 - 6 Implantação e monitoramento

Pré-processamento

■ Limpeza e tratamento de dados

Valores ausentes / Outliers / Ruídos

■ Transformação dos dados

Normalização / Padronização / Transformações polinomiais, logarítmicas, ...

■ Redução de dimensionalidade

Seleção de atributos / Transformação de atributos

■ Balanceamento dos dados

Oversampling / Undersampling

Ênfase do curso

■ Treinamento

- ▶ ~~Programar os principais algoritmos~~
- ▶ Como funciona os principais algoritmos
- ▶ Como usar os principais algoritmos
- ▶ Combinar diferentes modelos preditivos
- ▶ Otimizar os modelos gerados
- ▶ Gerar modelos automaticamente

■ Avaliação

- ▶ Procedimentos metodológicos
- ▶ Medidas de avaliação