



Aprendizado de máquina

O **aprendizado automático** ^(português brasileiro) ou a **aprendizagem automática** ^(português europeu) ou também **aprendizado de máquina** ^(português brasileiro) ou **aprendizagem de máquina** ^(português europeu) (em inglês: *machine learning*) é um subcampo da Engenharia e da ciência da computação que evoluiu do estudo de reconhecimento de padrões e da teoria do aprendizado computacional em inteligência artificial^[1]. Em 1959, Arthur Samuel definiu aprendizado de máquina como o "campo de estudo que dá aos computadores a habilidade de aprender sem serem explicitamente programados"^[2](livre tradução). O aprendizado automático explora o estudo e construção de algoritmos que podem aprender de seus erros e fazer previsões sobre dados^[3]. Tais algoritmos operam construindo um modelo a partir de inputs amostrais a fim de fazer previsões ou decisões guiadas pelos dados ao invés de simplesmente seguindo inflexíveis e estáticas instruções programadas. Enquanto que na inteligência artificial existem dois tipos de raciocínio (o indutivo, que extrai regras e padrões de grandes conjuntos de dados, e o dedutivo), o aprendizado de máquina só se preocupa com o indutivo.

Algumas partes do aprendizado automático estão intimamente ligadas (e muitas vezes sobrepostas) à estatística computacional; uma disciplina que foca em como fazer previsões através do uso de computadores, com pesquisas focando nas propriedades dos métodos estatísticos e sua complexidade computacional. Ela tem fortes laços com a otimização matemática, que produz métodos, teoria e domínios de aplicação para este campo. O aprendizado automático é usado em uma variedade de tarefas computacionais onde criar e programar algoritmos explícitos é impraticável. Exemplos de aplicações incluem filtragem de spam, reconhecimento ótico de caracteres (OCR)^[4], processamento de linguagem natural, motores de busca, diagnósticos médicos, bioinformática, reconhecimento de fala, reconhecimento de escrita, visão computacional e locomoção de robôs. O aprendizado de máquinas é às vezes confundido com mineração de dados^[5], que é um sub-campo que foca mais em análise exploratória de dados e é conhecido como aprendizado não supervisionado^{[6][7]}. No campo da análise de dados, o aprendizado de máquinas é um método usado para planejar modelos complexos e algoritmos que prestam-se para fazer predições- no uso comercial, isso é conhecido como análise preditiva. Esses modelos analíticos permitem que pesquisadores, cientistas de dados, engenheiros, e analistas possam "produzir decisões e resultados confiáveis e repetíveis" e descobrir os "insights escondidos" através do aprendizado das relações e tendências históricas nos dados.^[8]

As contribuições para a fundação da aprendizagem automática (aprendizado de máquina) feitas pelos cientistas John Hopfield e Geoffrey Hinton foram distinguidas em 2024 com o Prémio Nobel de Física. O Comité Nobel destacou os cientistas por "descobertas fundamentais e invenções que permitem aprendizagem automática com redes neurais artificiais".^{[9][10]}

Resumo

Tom M. Mitchell forneceu uma definição mais formal amplamente citada: "Diz-se que um programa de computador aprende pela experiência E , com respeito a algum tipo de tarefa T e performance P , se sua performance P nas tarefas em T , na forma medida por P , melhoram com a experiência E ."^[11] Esta definição das tarefas envolvidas no aprendizado de máquina é dada de forma fundamentalmente operacional, ao invés de cognitiva, seguindo a proposta de Alan Turing em seu artigo "Computadores e inteligência", em que a pergunta "As máquinas são capazes de pensar?" seja substituída pela pergunta "As máquinas são capazes de fazer o que (nós como entidades pensantes) podemos fazer?"^[12]

Tipos de problemas e tarefas

As tarefas de aprendizado de máquina são tipicamente classificadas em três categorias amplas, de acordo com a natureza do "sinal" ou "feedback" de aprendizado disponível para um sistema de aprendizado. Essas categorias são:^[13]

- *Aprendizado supervisionado*: São apresentadas ao computador exemplos de entradas e saídas desejadas, fornecidas por um "professor". O objetivo é aprender uma regra geral que mapeia as entradas para as saídas.
- *Aprendizado não supervisionado*: Nenhum tipo de etiqueta é dado ao algoritmo de aprendizado, deixando-o sozinho para encontrar estrutura nas entradas fornecidas. O aprendizado não supervisionado pode ser um objetivo em si mesmo (descobrir novos padrões nos dados) ou um meio para atingir um fim.
- *Aprendizado por reforço*: Um programa de computador interage com um ambiente dinâmico, em que o programa deve desempenhar determinado objetivo (por exemplo, dirigir um veículo). É fornecido, ao programa, feedback quanto a premiações e punições, na medida em que é navegado o espaço do problema. Outro exemplo de aprendizado por reforço é aprender a jogar um determinado jogo apenas jogando contra um oponente.

Entre o aprendizado supervisionado e o não supervisionado, está o aprendizado semi-supervisionado, em que o professor fornece um sinal de treinamento incompleto: um conjunto de dados de treinamento com algumas (muitas vezes várias) das saídas desejadas ausentes. A transdução é um caso especial deste princípio, em que o conjunto inteiro das instâncias do problema é conhecido no momento do aprendizado, mas com parte dos objetivos ausente.

Entre outras categorias de problemas de aprendizado de máquina, o meta-aprendizado aprende seu próprio viés indutivo com base em experiência prévias. A robótica epigenética, elaborada para o aprendizado de robôs, gera suas próprias sequências de situações de aprendizado (também chamadas de 'currículo'), para adquirir cumulativamente repertórios de novas habilidades, através de uma auto-exploração autônoma e da interação social com professores humanos. Faz uso de técnicas como aprendizado ativo, maturação, sinergias motoras e imitação.

Outra categorização de tarefas de aprendizado de máquina surge quando se considera a saída desejada em um sistema de aprendizado de máquina:

- Em classificação, entradas são divididas em duas ou mais classes, e o aprendiz deve produzir um modelo que vincula entradas não vistas a uma ou mais dessas classes (classificação multi-etiquetada). Isso é tipicamente abordado de forma supervisionada. A

filtragem de spam é um exemplo de classificação, em que as entradas são as mensagens de emails (ou outros) e as classes são "spam" ou "não spam".

- Em regressão, também um problema supervisionado, as saídas são contínuas, em vez de discretas.
- Em clustering, um conjunto de entradas é dividido em grupos. De maneira diferente da classificação, os grupos não são conhecidos previamente, tornando o clustering uma tarefa tipicamente não supervisionada^[14].
- A estimativa de densidades encontra a distribuição de entradas em algum espaço.
- A redução dimensional simplifica as entradas ao mapeá-las para um espaço de menor dimensão. A modelagem de tópicos é um problema relacionado, em que é fornecida ao programa uma lista de documentos em linguagem natural, solicitando que encontre documentos tratando de tópicos similares.

História e relação com outros campos

Como um esforço científico, o aprendizado de máquina cresceu a partir da busca pela inteligência artificial. Ainda nos princípios da IA como disciplina acadêmica, alguns pesquisadores já se interessavam em fazer máquinas aprenderem a partir de dados. Eles tentaram abordar o problema desde vários métodos simbólicos, assim como com o que foi então nomeado de "rede neural artificial"; estes eram majoritariamente perceptrons e outros modelos que mais tardes foram entendidos como reinvenções de modelos linear generalizados de estatística. A lógica probabilística também foi usada, especialmente em diagnósticos médicos automatizados.^{[13]:488}

No entanto, uma crescente ênfase na abordagem lógica, baseada em conhecimento, causou uma brecha entre a IA e o aprendizado de máquina. Sistemas probabilísticos eram atormentados por problemas teóricos e práticos de aquisição e representação de dados.^{[13]:488} Em 1980, sistemas especialistas haviam dominado a IA, e a estatística estava fora de uso.^[15] Trabalhos em aprendizado baseado em conhecimento/simbólico continuaram com a IA, levando até a programação lógica indutiva, mas agora as pesquisas mais voltadas para a estatística estavam fora do campo da IA, em reconhecimento de padrões e recuperação de informação.^{[13]:708–710; 755} Pesquisas em redes neuronais haviam sido abandonadas pela IA e ciência computacional em torno do mesmo período. Essa linha, também, foi continuada forma dos campos da IA e da ciência da computação, como "conexionismos", por pesquisadores de outras disciplinas incluindo Hopfield, Rumelhart e Hinton. Seu sucesso principal veio em meados da década de 1980 com a reinvenção da propagação para trás.^{[13]:25}

O aprendizado de máquina, reorganizado como um campo separado, começou a florescer na década de 1990. O campo mudou seu objetivo de alcançar a inteligência artificial para abordar problemas solucionáveis de natureza prática. Ele abandonou o foco em abordagens simbólicas que havia herdado da IA para métodos e modelos emprestados da estatística e da teoria da probabilidade.^[15] Ele também se beneficiou do crescente número de informação digitalizada disponível e a possibilidade de distribuir ela via internet.

O aprendizado de máquina e a mineração de dados com frequência fazem uso dos mesmos métodos e se sobrepõe significativamente, mas enquanto o aprendizado de máquina foca em fazer previsões, baseado em propriedades *conhecidas* aprendidas pelo dados de treinamento, a mineração de dados foca em descobrir as propriedades (previamente) *desconhecidas* nos dados (este é o passo dado na análise de extração de conhecimento na base de dados). A mineração de dados usa muitas métodos do aprendizado de máquina, mas com objetivos diferentes; por outro lado, o aprendizado de máquina também faz uso de

métodos da mineração de dados como "aprendizado não supervisionado" ou como um passo de processamento para melhorar a precisão do aprendiz. Muita da confusão entre essas duas comunidades de pesquisa (que com frequência tem conferências e periódicos separados, ECML PKDD sendo a grande exceção) vem da suposição básica com que eles trabalham: em aprendizado de máquina, a performance é normalmente avaliada com respeito a habilidade de reproduzir conhecimento *conhecido*, enquanto que com a extração de conhecimento e mineração de dados (KDD) a tarefa chave é o descobrimento de conhecimento previamente *desconhecido*. Avaliado com respeito ao conhecimento conhecido, um método uniforme (não supervisionado) será facilmente superado por outros métodos supervisionados, enquanto que em uma tarefa KDD típica, métodos supervisionados não podem ser usados devido a não disponibilidade de dados de treinamento.

O aprendizado de máquina também tem laços íntimos com a otimização: muitos dos problemas de aprendizado são formulados como minimização de algumas funções de perda em um conjunto exemplo de treinamentos. Funções de perda expressam a discrepância entre as previsões do modelo sendo treinado e as verdadeiras instâncias do problema (por exemplo, em classificação, o objetivo é designar etiquetas para instâncias, e modelos são treinados para prever corretamente as etiquetas previamente designadas de um conjunto exemplo). A diferença entre os dois campos surge do objetivo da generalização: enquanto que o algoritmo de otimização pode minimizar a perda em um conjunto de treinamento, o aprendizado de máquina está preocupado com a minimização da perda de amostras não vistas.^[16]

Relação com estatística

O aprendizado de máquina e a estatística são campos intimamente relacionados. De acordo com Michael I. Jordan, as ideias do aprendizado de máquina, dos princípios metodológicos às ferramentas teóricas, tem uma longa pré-história na estatística.^[17] Ele também sugeriu o termo ciência de dados como um substituto para chamar o campo como um todo.^[17]

Leo Breiman distinguiu dois paradigmas da modelagem estatística: modelo de dados e modelo algorítmico,^[18] onde "modelo algorítmico" significa mais ou menos os algoritmos do aprendizado de máquina como a Floresta aleatória.^[19]

Alguns estatísticos tem adotado métodos do aprendizado de máquinas, levando ao campo combinado que eles chamam de *aprendizado estatístico*.^[20]

Teoria

Um dos objetivos centrais de um aprendiz é generalizar a partir de suas experiências.^{[21][22]} Generalização neste contexto é a habilidade de uma máquina aprendiz de desempenhar com precisão em novos, não vistos, exemplos/tarefas depois de ter experimentado um conjunto de dados de aprendizado. Os exemplos de treinamento vem de algumas, geralmente desconhecidas, distribuições de probabilidade (consideradas representativas do espaço de ocorrência) e o aprendiz tem de construir um modelo geral sobre este espaço que o permita produzir previsões o suficientemente precisas em novos casos.

A análise computacional de algoritmos de aprendizado de máquina e sua performance é um ramo da ciência da computação teórica conhecida como teoria do aprendizado computacional. Porque os conjuntos de treinamento são finitos e o futuro é incerto, a teoria de aprendizado normalmente não guarda

garantias para a performance dos algoritmos. Em vez disso, previsões probabilísticas para a performance são bastante comuns. O trade-off entre variância e viés é uma forma de qualificar o erro de generalização.

Para uma melhor performance no contexto de generalização, a complexidade da hipótese deveria combinar com a complexidade da função subjacente aos dados. Se a hipótese é menos complexa que a função, então o modelo sub-ajustou (*underfitting*) os dados. Se a complexidade do modelo é aumentada em resposta, então o erro de treinamento diminui. Mas se a hipótese é muito complexa, então o modelo foi sobreajustado (*overfitting*), e a generalização será mais pobre.^[23]

Em adição aos limites da performance, teóricos do aprendizado computacional estudam a complexidade do tempo e a viabilidade do aprendizado. Na teoria do aprendizado computacional, uma computação é considerada viável se puder ser feita em tempo polinomial. Há dois tipos de resultados de complexidade temporal. Resultados positivos mostram que uma certa classe de funções pode ser aprendida em tempo polinomial. Resultados negativos mostram que certas classes não podem ser aprendidas em tempo polinomial.

Abordagens

Aprendizado/aprendizagem baseado em árvores de decisão

Aprendizado baseado em árvores de decisão usa a árvore de decisão como um modelo de previsão, o qual mapeia as observações sobre um item às conclusões sobre o valor do objetivo desse item.

Aprendizado/aprendizagem por regras de associação

Aprendizado por regras de associação é um método para descobrir relações interessantes entre variáveis em base de dados grandes.

Rede neural artificial

Um algoritmo de aprendizado de rede neural artificial (aprendizagem de rede neuronal artificial), normalmente chamado de "rede neural" ou "rede neuronal" (RN), é um algoritmo do tipo "machine learning" que é inspirado na estrutura e aspectos funcionais das redes neurais biológicas. Computações são estruturadas em termos de um grupo interconectado de neurônios artificiais, processando informação usando uma abordagem de conexionismo na computação. Redes neurais modernas são ferramentas de modelagem de dados estatísticos não lineares. Normalmente eles são usados para modelar relações complexas entre entradas e saídas, para encontrar padrões nos dados, ou para capturar a estrutura estatística em uma distribuição de probabilidade conjunta desconhecida entre variáveis observáveis.

Aprendizado profundo (aprendizagem profunda)

Os preços mais baixos do hardware e o desenvolvimento de GPUs para uso pessoal nos últimos anos contribuiu para o desenvolvimento do conceito de aprendizado profundo, que consiste em múltiplas camadas escondidas em uma rede neural artificial. Esta abordagem tenta modelar a forma com que o cérebro humano processa luz e som na visão e escuta. Alguns aplicações úteis do aprendizado profundo são visão computacional e reconhecimento de fala.^[24]

Lógica de programação indutiva

Lógica de programação indutiva (LPI) é uma abordagem que regra o aprendizado fazendo uso de programação lógica como uma representação uniforme para exemplos de inputs, conhecimento de pano de fundo, e hipóteses. Dada uma codificação do pano de fundo conhecido do conhecimento e um conjunto de exemplos representados como uma base de dados lógica de fatos, um sistema LPI derivará uma lógica hipotética que envolve todos os exemplos positivos e não negativos. A programação indutiva é um campo relacionado que considera qualquer tipo de linguagem de programação para representar hipóteses (e não apenas programações lógicas), tais como as programações funcionais.

Máquinas de vetores de suporte

Máquinas de vetores de suporte (MVSs) são um conjunto relacionado de métodos de aprendizado supervisionado usados para classificação e regressão. Dado um conjunto de exemplos de treinamento, cada um marcado como pertencente de uma ou duas categorias, um algoritmo de treino SVM constrói um modelo que prediz se um novo exemplo cai dentro de uma categoria ou outra.

Clustering

Análise de clusters é a atribuição de um conjunto de observações à subconjuntos (chamados clusters) de forma que as observações dentro de um mesmo cluster são similares de acordo com algum critério ou critérios pré-designados, enquanto que observações feitas em clusters diferentes não são similares. Diferentes técnicas de clustering fazem diferentes suposições sobre a estrutura dos dados, freqüentemente definida por algumas *métricas de similaridade* e avaliados, por exemplo, por *compacidade interna* (similaridade entre membros de um mesmo cluster) e *separação* entre clusters diferentes. Outros métodos são baseado em *estimações de densidade* e *gráficos de conectividade*. Clustering é um método de aprendizado não supervisionado e uma técnica comum em análise de dados estatísticos.

Redes Bayesianas

Uma rede bayesiana, rede de opinião ou um modelo gráfico acíclico dirigido é um modelo gráfico probabilístico que representa um conjunto de variáveis aleatórias e suas independências condicionais via um grafo acíclico dirigido (GAD). Por exemplo, uma rede bayesiana poderia representar as relações probabilísticas entre doenças e sintomas. Dado um sintoma, a rede pode ser usada para computar as probabilidades da presença de várias doenças. Existem algoritmos eficientes que desempenham inferência e aprendizado.

Aprendizado por reforço

O aprendizado por reforço se preocupa com o como um *agente* deve *agir* em um *ambiente* de forma que maximize alguma noção de *recompensa* a longo tempo. Os algoritmos de aprendizado por reforço tentam encontrar a *política* que mapeia os estados do mundo às ações que o agente deve ter nesses estados. Aprendizado por reforço se distingue do problema do aprendizado supervisionado no sentido em que pares de input/output corretos nunca são apresentados, nem as ações sub-ótimas são explicitamente corrigidas.

Aprendizado por representação

Vários algoritmos de aprendizado, a maioria algoritmos de aprendizado não supervisionado, tem como objetivo descobrir melhores representações dos inputs que são dados durante o treinamento. Exemplos clássicos incluem análise de componentes principais e análise de clusters. Os algoritmos de aprendizado por representação com frequência tentam preservar a informação e seu input mas transformando-a de forma que a torne útil, frequentemente como um passo pré-processamento antes de desempenhar classificações ou previsões, permitindo a reconstrução dos inputs vindos de dados geradores de distribuição desconhecidos, enquanto não sendo necessariamente fiel à configurações que são implausíveis sob essa distribuição.

Algoritmos de aprendizado múltiplo tentam fazer isso sob a restrição de que a representação aprendida é de baixa dimensão. Algoritmos de código esperso tentam fazer isso sob a restrição de que a representação aprendida é espersa (tem muitos zeros). O algoritmo de aprendizado em subespaço multilinear tem como objetivo aprender representações de baixa dimensão diretamente das representações de tensores para dados multidimensionais, sem os transformar em vetores de alta dimensão.^[25] Algoritmos de aprendizado profundo descobrem múltiplos níveis de representação, ou uma hierarquia de características, com um nível mais alto, características mais abstratas definidas em termos de (ou geradas a partir de) características de nível mais baixo. Tem sido argumentado que uma máquina inteligente é aquela que aprende uma representação que desembaraça os fatores subjacentes de variação que explicam os dados observados.^[26]

Aprendizado por similaridade e métrica

Neste problema, se dá a máquina aprendiz pares de exemplos que são considerados similares e pares de objetos menos similares. Então ela precisa aprender uma função de similaridade (ou uma função de distancia métrica) que possa prever se novos objetos são similares. Isso é às vezes usado em sistemas de recomendação.

Aprendizado por dicionário esperso

Neste método, um dado é representado como uma combinação de ligar funções bases, e os coeficientes são assumidos como espersos. Deixe x ser um dado d -dimensional, D ser um d por n matrizes, onde cada coluna de D representa uma função base. r é o coeficiente para representar x usando D . Matematicamente, aprendizado por dicionário esperso significa resolver $x \approx Dr$ onde r é esperso. Falando genericamente, se assume que n é maior que d para permitir a liberdade para uma representação espersa.

Aprender um dicionário junto com representações espersas é fortemente NP-completo e também difícil de resolver aproximadamente.^[27] Um método heurístico popular para aprendizado por dicionário esperso é o K-SVD.

Aprendizado por dicionário esperso tem sido aplicado a vários contextos. Em classificação, o problema é determinar a quais classes um dado previamente não visto pertence. Suponha que um dicionário para cada classe já tenha sido construído. Então um novo dado é associado com a classe de forma que esteja o melhor espersalmente representado pelo dicionário correspondente. Aprendizado por dicionário esperso também tem sido aplicado a suavização de imagem. A ideia chave é que um pedaço de imagem limpa pode ser representado espersalmente por um dicionário de imagem, mas o ruído não.^[28]

Detecção de anomalias

Em mineração de dados, a **detecção de anomalias** (também **detecção de outliers** ^[29]) é a identificação de itens, eventos ou observações raros que levantam suspeitas por serem significativamente diferentes da maioria dos dados. ^[1] Normalmente, os itens anômalos se referem a algum tipo de problema, como fraude bancária, defeito estrutural, problemas médicos ou erros em um texto. Anomalias são também referidas como outliers, ruído, desvios e exceções. ^[30]

Algoritmos genéticos

Um algoritmo genético (AG) é uma busca heurística que imita o processo de seleção natural e usa métodos com mutação e recombinação para gerar novos genótipos na esperança de encontrar boas soluções para um dado problema. Em aprendizado de máquinas, algoritmos genéticos encontraram alguma utilidade em 1980 e 1990. ^{[31][32]} Vice versa, técnicas do aprendizado de máquina tem sido usadas para melhorar a performance de algoritmos genéticos e evolutivos. ^[33]

Aplicações

Aplicações para o aprendizado de máquina incluem:

- Websites adaptativos
- Computação afetiva
- Computação quântica ^[34]
- Bioinformática
- Interface cérebro-computador
- Quimioinformática
- Classificação de sequências de DNA
- Anatomia computacional
- Visão computacional, incluindo reconhecimento de objetos
- Detecção de fraude de cartão de crédito
- Jogos de estratégia
- Recuperação de informação
- Detecção de fraude virtual
- Marketing
- Percepção de máquina
- Diagnósticos médico
- Economia
- Processamento de linguagem natural
- Entendimento de linguagem natural
- matemática e meta-heurística
- Publicidade online
- Sistema de recomendação
- Locomoção de robôs
- Mecanismos de busca
- Análise de sentimento ou mineração de opinião
- Mineração de padrões sequenciais
- Engenharia de software
- Reconhecimento de fala e reconhecimento de escrita
- Análises no mercado de ações
- Monitoramento estrutural da saúde
- Reconhecimento de padrões sintáticos

Em 2006, a companhia de filmes online Netflix fez a primeira competição "Netflix Prize" para encontrar um programa que melhor prediria as performances dos usuários e melhoraria a precisão do algoritmo de recomendação Cinematch existente em ao menos 10%. Um time composto por pesquisadores da AT&T Labs em colaboração com o time Big Chaos e Pragmatic Theory construíram um modelo conjunto para ganhar o grande prêmio em 2009 de 1 milhão de dólares. Logo após o prêmio ser concedido, a Netflix se deu conta que as avaliações dos usuários não eram as melhores indicações de seus padrões de filmes e séries vistos ("tudo é recomendação") e mudaram seu motor de recomendação. ^[35]

Em 2010 o Periódico do Wall Street escreveu sobre a firma de gestão de dinheiro Rebellion Research que usava o aprendizado de máquina para prever os movimentos econômicos. O artigo descrevia a previsão da Rebellion Research sobre a crise financeira e a recuperação econômica.^[36]

Em 2014 foi relatado que um algoritmo de aprendizado de máquina foi aplicado em Historia da Arte para estudar as pinturas de arte e que ele pode ter revelado influências entre artistas previamente não reconhecidas.^[37]

Em 2015, o serviço de streaming de música Spotify criou a playlist chamada Descobertas da Semana, que funciona como uma curadoria digital. O algoritmo responsável por esta playlist utiliza técnicas de Filtragem Colaborativa, Processamento de Linguagem Natural e Processamento de Sinais de Áudio através de Redes Neurais Convolucionais para compor a playlist semanalmente.^[38]

Avaliação de modelos

Modelos de classificação de aprendizado de máquina podem ser validados por técnicas como validação cruzada, onde os dados são divididos em conjuntos de teste e treinamento e medidas apropriadas como precisão são calculadas para ambos os conjuntos de dados e comparadas. Para além da precisão, sensibilidade (Avaliação Positiva Verdadeira) e especificidade (Avaliação Negativa Verdadeira) podem prover modos de modelos de avaliação. De forma similar, Avaliações Positivas Falsas assim como Avaliações Negativas Falsas poder ser computadas. Curvas Receptoras de Operação (CRO) em conjunto com a Área em baixo da CRO (AUC) oferecem ferramentas adicionais para a classificação de modelos de avaliação. Graus maiores de AUC estão associados a um modelo de melhor performance.^[39]

Software

Software suites contendo uma variedade de algoritmos de aprendizado de máquina incluem os seguintes:

Software de fonte aberta e gratuita

- | | |
|---------------------------------|-----------------------|
| ■ caret | ■ <u>OpenAI</u> |
| ■ dlib | ■ <u>OpenCV</u> |
| ■ ELKI | ■ <u>OpenNN</u> |
| ■ Encog | ■ Orange |
| ■ <u>GNU Octave</u> | ■ R |
| ■ H2O | ■ <u>scikit-learn</u> |
| ■ Mahout | ■ scikit-image |
| ■ Mallet (software project) | ■ Shogun |
| ■ mlpy | ■ <u>TensorFlow</u> |
| ■ MLPACK | ■ Torch |
| ■ MOA (Massive Online Analysis) | ■ Spark |
| ■ ND4J with Deeplearning4j | ■ Yooreeka |
| ■ NuPIC | ■ Weka |

Software particular com edições de fonte aberta e gratuita

- KNIME
- RapidMiner

Software proprietário

- Amazon Machine Learning
- Angoss KnowledgeSTUDIO
- Ayasdi
- BigML
- Databricks
- Google APIs|Google Prediction API
- IBM SPSS Modeler
- KXEN Modeler
- LIONSolve
- Mathematica
- MATLAB
- Microsoft Azure Machine Learning
- Neural Designer
- NeuroSolutions
- Oracle Data Mining
- RCASE
- SAS Enterprise Miner
- Splunk
- STATISTICA Data Miner

Veja também esta lista de pacotes em muitas linguagens de programação: Awesome Machine Learning (<https://github.com/josephmisiti/awesome-machine-learning#awesome-machine-learning->).

Periódicos

- *Journal of Machine Learning Research*
- *Machine Learning*
- *Neural Computation*
- *International Journal of Machine Learning and Cybernetics*

Conferências

- Conference on Neural Information Processing Systems
- International Conference on Machine Learning

Ver também

- Aprendizagem profunda
- Aprendizado por transferência
- Inteligência artificial
- Lista de projetos de inteligência artificial
- Redes Neurais
- Clustering
- Bioinformática
- Visão computacional
- Mineração de dados
- Reconhecimento de padrões
- Máquina de Vetores de Suporte

- Big data
- Ciência cognitiva
- Inteligência computacional
- Neurociência computacional


Ligações externas

- «Conceitos sobre Aprendizado de Máquina» (<http://dcm.ffclrp.usp.br/~augusto/publications/2003-sistemas-inteligentes-cap4.pdf>) (PDF)
- Curso online no Coursera (<https://www.coursera.org/course/ml>), ministrado por Andrew Ng (<http://www.andrewng.org>). Utiliza GNU Octave. O curso é uma versão livre do curso da Stanford University ministrado por Ng, cujas aulas também estão disponíveis gratuitamente (<https://see.stanford.edu/Course/CS229>).
- Aula sobre Aprendizagem por Reforço (<http://professor.ufabc.edu.br/~ronaldo.prati/InteligenciaArtificial/reinforcement-learning.pdf>), ministrada pelo professor Ronaldo C. Prati, Professor Adjunto da UFABC

Referências

1. «Definição: machine learning» (<https://www.britannica.com/technology/machine-learning>). *Enciclopedia Britannica* (em inglês)
2. Simon, Phil (2013). *Too Big to Ignore: The Business Case for Big Data* (https://books.google.com.br/books?id=Dn-Gdoh66sgC&pg=PA89&redir_esc=y#v=onepage&q&f=false). [S.l.]: Wiley. 89 páginas. ISBN 978-1-118-63817-0
3. Kohavi, Ron; Provost, Foster (1998). «Glossary of terms» (<http://ai.stanford.edu/~ronnyk/glossary.html>). *Machine Learning* 30: 271–274
4. Wernick, Yang, Brankov, Yourganov and Strother (Julho, 2010). «Machine Learning in Medical Imaging». *IEEE Signal Processing Magazine*. **27** (4): 25-38
5. Mannila, Heikki (1996). «Data mining: machine learning, statistics, and databases». *Int'l Conf. Scientific e Statistical Database Management. IEEE Computer Society*
6. Friedman, Jerome H. (1998). «Data Mining and Statistics: What's the connection?». *Computing Science and Statistics* 29 (1): 3–9
7. «O que é Machine Learning?» (<https://openimob.com/blog/index.php/2018/08/13/o-que-e-machine-learning/>)
8. «Machine Learning: What it is and why it matters» (http://www.sas.com/it_it/insights/analytic/machine-learning.html) (em inglês). SAS The Power to Know. Consultado em 10 de outubro de 2016
9. Comité Nobel. «The Nobel Prize in Physics 2024 - press release» (<https://www.nobelprize.org/prizes/physics/2024/press-release/>). *nobel.se* (em inglês). 8-10-2024. Consultado em 8 de outubro de 2024
10. publico.pt. «As bases para a inteligência artificial vencem o Prémio Nobel da Física de 2024» (<https://www.publico.pt/2024/10/08/ciencia/noticia/bases-inteligencia-artificial-vencem-premio-nobel-fisica-2024-2106946>). 8-10-2024. Consultado em 8 de outubro de 2024
11. M. Mitchell, Tom (1997). *Machine Learning*. [S.l.]: McGraw Hill. 2 páginas. ISBN 0-07-042807-7

12. Harnad, Steven (2008). *The Annotation Game: On Turing (1950) on Computing, Machinery, and Intelligence* (<http://eprints.soton.ac.uk/262954/>). in Epstein, Robert; Peters, Grace: *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer* (em inglês). [S.l.]: Springer. pp. 23–66. Consultado em 10 de outubro de 2016
13. Russell, Stuart; Norvig, Peter (2003). *Artificial Intelligence: A Modern Approach* (https://drive.google.com/file/d/0ByctuzDL9IM9ZGFhNGQ1NjgtODAwOS00NDAYLWE3MDgtZWJjZTRhNDE4MGY5/view?hl=en_GB) (em inglês) 2 ed. [S.l.]: Prentice Hall. ISBN 978-0137903955
14. «Teachable Machine» (<https://teachablemachine.withgoogle.com/>). *teachablemachine.withgoogle.com*. Consultado em 7 de maio de 2021
15. Langley, Pat (2011). «The changing science of machine learning» (<http://link.springer.com/article/10.1007%2Fs10994-011-5242-y>). *Machine Learning* (em inglês). **82** (3): 275–279. doi:10.1007/s10994-011-5242-y (<https://dx.doi.org/10.1007%2Fs10994-011-5242-y>). Consultado em 10 de outubro de 2016
16. Le Roux, Nicolas; Bengio, Yoshua; Fitzgibbon, Andrew (2012). *Improving First and Second-Order Methods by Modeling Uncertainty In Sra, Suvrit; Nowozin, Sebastian; Wright, Stephen J. Optimization for Machine Learning* (em inglês). [S.l.]: MIT Press. 404 páginas
17. I. Jordan, Michael (9 de outubro de 2014). «Statistics and machine learning» (https://www.reddit.com/r/MachineLearning/comments/2fxi6v/ama_michael_i_jordan/ckelmtt?context=3) (em inglês). Reddit. Consultado em 11 de outubro de 2016
18. Breiman, Leo (2001). «Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author)» (<http://projecteuclid.org/euclid.ss/1009213726>). The Institute of Mathematical Statistics. *Statistical Science* (em inglês). **16** (3): 199-231
19. Izbicki, Santos (2018). «Machine Learning sob a ótica estatística» (<http://www.rizbicki.ufscar.br/sml/>). Consultado em 23 de agosto de 2018
20. Gareth, James; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert (2013). *An Introduction to Statistical Learning* (<http://www-bcf.usc.edu/~gareth/ISL/>) (em inglês). [S.l.]: Springer. 7 páginas
21. Bishop, Christopher (2006). *Pattern Recognition and Machine Learning* (em inglês). [S.l.]: Springer. ISBN 0-387-31073-8
22. Mehryar, Mohri; Rostamizadeh, Afshin; Talwalkar, Ameet (2012). *Foundations of Machine Learning* (em inglês). [S.l.]: MIT Press. ISBN 978-0-262-01825-8
23. Alpaydin, Ethem (2010). *Introduction to Machine Learning* (https://books.google.com/books?id=NP5bBAAQBAJ&printsec=frontcover&dq=ethem+alpaydin&hl=tr&sa=X&redir_esc=y#v=onepage&q=ethem%20alpaydin&f=false) (em inglês). [S.l.]: MIT Press
24. Honglak Lee, Roger Grosse, Rajesh Ranganath, Andrew Y. Ng. (2009). «Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations» (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.149.802&rep=rep1&type=pdf>). Proceedings of the 26th Annual International Conference on Machine Learning (em inglês)
25. Lu, Haiping; Plataniotis, K.N.; Venetsanopoulos, A.N. (2011). «A Survey of Multilinear Subspace Learning for Tensor Data» (http://www.dsp.utoronto.ca/~haiping/Publication/SurveyMSL_PR2011.pdf) (PDF). *Pattern Recognition* (em inglês). **44** (7): 1540–1551. doi:10.1016/j.patcog.2011.01.004 (<https://dx.doi.org/10.1016%2Fj.patcog.2011.01.004>). Consultado em 17 de outubro de 2016
26. Bengio, Yoshua (2009). *Learning Deep Architectures for AI* (https://books.google.com.br/books?id=cq5ewg7FniMC&pg=PA3&redir_esc=y#v=onepage&q&f=false). [S.l.]: Now Publishers Inc. pp. 1–3. ISBN 978-1-60198-294-0
27. Tillmann, Andreas M. (2015). «On the Computational Intractability of Exact and Approximate Dictionary Learning». *IEEE Signal Processing Letters*. **22** (1): 45-49
28. Aharon, Michael; M. Elad; A. Bruckstein (2006). «K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation». *IEEE Transactions on Signal Processing*. **54** (11): 4311-4322

29. Zimek, Arthur; Schubert, Erich (2017), «Outlier Detection», ISBN 9781489979933, Springer New York, *Encyclopedia of Database Systems*: 1–5, doi:10.1007/978-1-4899-7993-3_80719-1 (https://dx.doi.org/10.1007%2F978-1-4899-7993-3_80719-1)
30. «A Survey of Outlier Detection Methodologies» (<http://eprints.whiterose.ac.uk/767/1/hodgvej4.pdf>) (PDF). *Artificial Intelligence Review*. **22**. CiteSeerX 10.1.1.318.4023 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.318.4023>)  doi:10.1007/s10462-004-4304-y (<https://dx.doi.org/10.1007%2Fs10462-004-4304-y>)
31. Goldberg, David E; Holland, John H. (1988). «Genetic algorithms and machine learning». *Machine Learning* (em inglês). **3** (2): 95–99. doi:10.1007/bf00113892 (<https://dx.doi.org/10.1007%2Fbf00113892>)
32. Michie, Donald; Spiegelhalter, David; C.C., Taylor (1994). *Machine Learning, Neural and Statistical Classification* (<http://www1.maths.leeds.ac.uk/~charles/statlog/>) (em inglês). [S.l.]: Ellis Horwood
33. Zhang, Jun; Zhan, Zhi-hui; Lin, Ying; Chen, Ni; Gong, Yue-jiao; Zhong, Jing-hui; Chung, Henry S.H.; Li, Yun; Shi, Yu-hui (2011). «Evolutionary Computation Meets Machine Learning: A Survey» (<http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6052374>). *Computational Intelligence Magazine* (em inglês). **6** (4): 68-75. doi:10.1109/mci.2011.942584 (<https://dx.doi.org/10.1109%2Fmci.2011.942584>)
34. «Improving quantum computation with classical machine learning» (<https://www.techexplorist.com/improving-quantum-computation-classical-machine-learning/26856/>). *Tech Explorist* (em inglês). 4 de outubro de 2019. Consultado em 7 de outubro de 2019
35. «The Netflix Tech Blog: Netflix Recommendations: Beyond the 5 stars (Part 1)» (<http://techblog.netflix.com/2012/04/netflix-recommendations-beyond-5-stars.html>). Consultado em 17 de outubro de 2016
36. «Artificial Intelligence' Gains Fans Among Investors - WSJ» (<http://www.wsj.com/news/articles/SB10001424052748703834604575365310813948080>) (em inglês). Consultado em 17 de outubro de 2016
37. «When A Machine Learning Algorithm Studied Fine Art Paintings, It Saw Things Art Historians Had Never Noticed» (<https://medium.com/the-physics-arxiv-blog/when-a-machine-learning-algorithm-studied-fine-art-paintings-it-saw-things-art-historians-had-never-b8e4e7bf7d3e#.xsd2hk26i>) (em inglês). Consultado em 17 de outubro de 2016
38. Ciocca, Sophia. «How Does Spotify Know You So Well? – Member Feature Stories» (<https://medium.com/s/story/spotify-discover-weekly-how-machine-learning-finds-your-new-music-19a41ab76efe>). *Medium*. Consultado em 3 de dezembro de 2018
39. Catal, Cagatay (2012). «Performance Evaluation Metrics for Software Fault Prediction Studies» (http://www.uni-obuda.hu/journal/Catal_36.pdf) (PDF). *Acta Polytechnica Hungarica* (em inglês). **9** (4)

Para mais leituras

- Brocardo ML, Traore I, Woungang I, Obaidat MS. "Authorship verification using deep belief network systems (<http://onlinelibrary.wiley.com/doi/10.1002/dac.3259/full>)". *Int J Commun Syst*. 2017. doi:10.1002/dac.3259
- Trevor Hasti, Robert Tibshirani e Jerome H. Friedman (2001). *The Elements of Statistical Learning* (<https://web.archive.org/web/20091110212529/http://www-stat.stanford.edu/~tibs/ElemStatLearn/>), Springer. ISBN 0-387-95284-5.
- Pedro Domingos (September 2015), *The Master Algorithm*, Basic Books, ISBN 978-0-465-06570-7
- Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar (2012). *Foundations of Machine Learning* (<http://www.cs.nyu.edu/~mohri/mlbook/>), The MIT Press. ISBN 978-0-262-01825-8.

- Ian H. Witten and Eibe Frank (2011). *Data Mining: Practical machine learning tools and techniques* Morgan Kaufmann, 664pp., ISBN 978-0-12-374856-0.
- David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms* (<http://www.inference.phy.cam.ac.uk/mackay/itila/book.html>) Cambridge: Cambridge University Press, 2003. ISBN 0-521-64298-1
- Richard O. Duda, Peter E. Hart, David G. Stork (2001) *Pattern classification* (2nd edition), Wiley, New York, ISBN 0-471-05669-3.
- Christopher Bishop (1995). *Neural Networks for Pattern Recognition*, Oxford University Press. ISBN 0-19-853864-2.
- Vladimir Vapnik (1998). *Statistical Learning Theory*. Wiley-Interscience, ISBN 0-471-03003-1.
- Ray Solomonoff, *An Inductive Inference Machine*, IRE Convention Record, Section on Information Theory, Part 2, pp., 56-62, 1957.
- Ray Solomonoff, "An Inductive Inference Machine (<http://world.std.com/~rjs/indinf56.pdf>)" Um relato privado que circulou na Dartmouth Summer Research Conference sobre IA em 1956.

Obtida de "https://pt.wikipedia.org/w/index.php?title=Aprendizado_de_máquina&oldid=68787599"