

MÓDULO I

INTRODUÇÃO AOS FUNDAMENTOS DA INTELIGÊNCIA ARTIFICIAL

Autores:

ALEX REIMANN CUNHA LIMA

ANA JULIA BRITES ALVES

VINICIUS MANHÃES GABRIEL DE BRITO CAVALCANTI

Ilustrações:

DALL-E STABLE DIFFUSION

CAMPUS ITABORAÍ 2024



IDENTIFICAÇÃO INSTITUCIONAL

IFFLUMINENSE — *Campus*: Itaboraí

CNPJ: 10.779.511/0001-07

Endereço completo: Rua Izaura Pantoja, 167-333, bairro Nova Cidade, Itaboraí/ RJ

Fone/Fax de contato: (22) 2737-5624 (gabinete da reitoria)

E-mail de contato: campus.itaborai@iff.edu.br

Diretor Geral: Vicente de Paulo Santos de Oliveira





MINISTÉRIO DA EDUCAÇÃO SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA FLUMINENSE CAMPUS ITABORAÍ

REITOR

JEFFERSON MANHÃES DE AZEVEDO

PRÓ-REITOR DE ENSINO

CARLOS ARTUR DE C. ARÊAS

DIRETOR GERAL DO CAMPUS ITABORAÍ

VICENTE DE PAULO SANTOS DE OLIVEIRA

DIRETOR DE ENSINO

DANIEL PINHEIRO CAETANO DAMASCENO

COORDENADOR DE PESQUISA E EXTENSÃO

WANDERSON AMARAL DA SILVA

COORDENADOR DE EDUCAÇÃO A DISTÂNCIA

RENATO MEIRA DE SOUSA DUTRA

COORDENADORES DO CURSO

ALEX REIMANN CUNHA LIMA FABIO DA SILVEIRA SEVERIANO GERALDO FURTADO NETO

COMISSÃO RESPONSÁVEL PELA ELABORAÇÃO DO PROJETO PEDAGÓGICO DO CURSO

ALEX REIMANN CUNHA LIMA
ANA JULIA BRITES ALVES
FABIO DA SILVEIRA SEVERIANO
GERALDO FURTADO NETO
RENATO MEIRA DE SOUSA DUTRA
VINICIUS MANHÃES GABRIEL DE BRITO CAVALCANTI
WANDERSON AMARAL DA SILVA



SUMÁRIO

CON	NCEITOS DE APRENDIZADO DE MÁQUINA	5
2.1	Um pouco de Matemática e alguns termos	6
2.2	Introdução ao Aprendizado de Máquina	7
2.3	O que torna o Aprendizado de Máquina especial?	7
2.4	Como funciona o Aprendizado de Máquina?	8
2.5	Aplicações do Aprendizado de Máquina	9
2.6	Tipos de Aprendizado	11 12 13 13 14
2.7	Separação dos dados	17 17



2 CONCEITOS DE APRENDIZADO DE MÁQUINA

Bem-vindo(a) ao Capítulo 2! Você já se perguntou como um serviço de streaming consegue recomendar novos filmes ou seriados de acordo com seus gostos? Parece mágica, não é mesmo? Bom, vou lhe contar um segredo: Não é mágica, é algo chamado de AM (Aprendizado de Máquina), uma área muito importante da IA.

Imagine como seria se você pudesse ensinar ao seu computador a aprender com as experiências dele, assim como você aprende com as suas. AM é isso! É como dar ao computador um cérebro que fica mais inteligente ao longo do tempo. Neste capítulo, vamos mergulhar no mundo fantástico do AM. Vamos começar entendendo o que ele é de fato, e então vamos explorar as diferentes maneiras dos computadores aprenderem. Alguns aprendem seguindo instruções, outros aprendem observando e adivinhando, e outros até mesmo conseguem aprender por meio de jogos, tentando a melhor estratégia para vencê-los.

Vamos também dar uma olhada nas receitas, ou **algoritmos**, que guiam os computadores em seu aprendizado. Esses algoritmos atuam como receitas de bolo para o cérebro do seu computador, dizendo a ele como entender as informações e aprender com elas. Ao final deste capítulo, você vai aprender como essas tecnologias fabulosas estão criando um futuro no qual os computadores não são apenas ferramentas, mas também parceiros na solução de problemas e na realização de novas descobertas.

Portanto, aperte o cinto e prepare-se para uma incrível jornada no mundo do AM. Será uma viagem fantástica!

algoritmo Um como uma receita para resolver um É um problema. conjunto de instruções (passo-a-passo) que diz a você exatamente o que fazer para resolver um problema ou alcançar um objetivo. Assim como uma receita pode lhe ajudar a fazer um bolo dizendo ingredientes auais devem ser usados e que passos devem ser seguidos, um algoritmo guia você pela solução de problemas tais como desafios matemáticos, ordenação de números determinação menor caminho até a escola. É um modo sistemático de resolução de problemas que tanto seres humanos quanto computadores usam.



2.1 Um pouco de Matemática e alguns termos

Sabemos agora sobre as coisas incríveis que a IA é capaz de fazer. Mas você deve estar se perguntando: o que, de fato, a IA faz na prática? Vamos entender isso com um exemplo bem simples. Imagine alguém que vende salgadinhos. Essa pessoa alugou um espaço para vender seus salgadinhos a R\$7,00 cada e tem um custo fixo mensal de R\$500. Como saber quanto ela vai ganhar no fim do mês? Se cada salgadinho é vendido a R\$7,00 e o custo fixo é de R\$500,00, podemos usar uma conta simples para descobrir: multiplicamos o número de salgadinhos vendidos por 7 e depois subtraímos os R\$500,00 do custo fixo. Matematicamente, representamos isso assim:

$$f(x) = 7x - 500,$$

onde f(x) é o saldo mensal (que pode ser negativo no prejuízo ou positivo no lucro) e x é quantidade de salgados vendida (que pode ser variável).

Mas e se, em vez de apenas o custo fixo e o preço dos salgadinhos, houvesse cem coisas diferentes afetando quanto dinheiro se ganha (custos com funcionários, publicidade, demanda sazonal, mudanças no gosto dos consumidores etc), e não soubéssemos como cada uma delas influencia o resultado? Seria super difícil e demorado para uma pessoa fazer essa conta na mão. Porém, uma IA que utiliza AM pode fazer isso rapidamente e com muita precisão. A IA ajuda a resolver muitos problemas encontrando a melhor forma de fazer essas contas complicadas que explicam situações do dia a dia.

Essas contas, ou "funções", podem variar muito. Podem ser



desde calcular o caminho mais curto entre dois lugares (como o Google Maps faz) até sugerir um filme que você pode gostar, baseado no que você já assistiu. Quando usamos essas técnicas de IA para tentar prever coisas novas baseadas em informações que já temos, chamamos isso de **modelo**. Os detalhes que mudam o resultado (as coisas que afetam lucro final, no nosso exemplo) são chamados de **variáveis independentes**, e o resultado final (o lucro ou prejuízo, no exemplo) é a **variável dependente**.

2.2 Introdução ao Aprendizado de Máquina

O AM é uma importante categoria de IA que, à primeira vista, se parece como algo saído direto das histórias de ficção científica, mas que, de fato, já faz parte das nossas vidas. Você pode imaginar o AM como um processo para ensinar os computadores a aprender a tomar decisões por si próprios, de forma similar àquela que nós, seres humanos, utilizamos para aprender com nossas experiências. A diferença é que o AM utiliza **dados** e **algoritmos** no lugar de livros e aulas!

2.3 O que torna o Aprendizado de Máquina especial?

Imagine que você deseja ensinar uma amigo alienígena a distinguir maçãs de bananas. Você mostraria ele exemplos de cada um e apontaria (e explicaria) as diferenças. No mundo do AM, em vez de ensinar as regras de forma direta, nós alimentamos os computadores com quantidades grandes de imagens de banas e maçãs, e os algoritmos "aprendem" a distingui-las por conta própria.





Figura 1: Ensino de distinção entre maçãs e bananas para um modelo de IA.

2.4 Como funciona o Aprendizado de Máquina?

O processo se inicia sempre com os dados. Quanto mais dados fornecermos ao algoritmo, melhor será o aprendizado do computador. Esses dados são divididos em dois grupos: aqueles que usamos para **treinar** o algoritmo (conjunto de treino) e aqueles que usamos para **testar** sua precisão (conjunto de teste).

Assim que tivermos os dados, poderemos escolher um dentre vários algoritmos de AM. Você pode imaginar cada um desses algoritmos como uma diferente receita (passo-a-passo) que o computador irá seguir para tentar aprender com os dados. Existem muitos tipos diferentes de receitas, e a escolha depende de o que desejamos que o computador aprenda.

Após o treinamento, nós devemos testar nosso modelo utilizando o conjunto de dados de teste para verificar o que foi aprendido. Se o desempenho não for satisfatório, podemos ajustar nosso algoritmo ou usar uma quantidade maior de dados e tentar novamente.



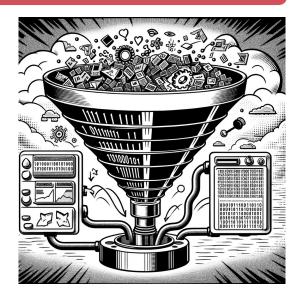


Figura 2: Processo de aprendizado de máquina, no qual dados são alimentados em um algoritmo, que então faz previsões ou toma decisões.

Esse processo é iterativo, ou seja, nós devemos repetir esse ciclo várias vezes até obtermos o resultado desejado.

2.5 Aplicações do Aprendizado de Máquina

O AM já está em todos os lugares! Sempre que a Netflix recomenda seriados de que você gosta, você já está presenciando o AM em ação. Ou ainda quando o Google Fotos reconhece seu rosto em fotos, você também está diante de um exemplo de AM.

2.6 Tipos de Aprendizado

Já vimos que, assim como temos a habilidade de aprender novos conceitos para uma prova, as máquinas (computadores) podem também ser ensinados a aprender com os dados. Nesta seção, veremos os três principais tipos de aprendizado: supervisado, não-supervisionado e reforçado.





Figura 3: Exemplos de aplicação de aprendizado de máquina: recomendações do Netflix e reconhecimento de rosto no Google Fotos.

Imagine que você está aprendendo a andar de bicicleta. Seus pais podem lhe acompanhar segurando a bicicleta até que você consiga descobrir como manter o equilíbrio. Essa é mais ou menos a ideia do **aprendizado supervisionado**, no qual um professor ou "supervisor" fornece à máquina **exemplos rotulados**.

Mas e se não houvesse ninguém para segurar sua bicicleta? Você pode tentar estratégias diferentes, tais como jogar o peso para um lado ou outro, ou ainda pedalar mais rápido, até descobrir o que funciona melhor. Isso é similar ao **aprendizado não-supervisionado**, no qual a máquina tem de encontrar padrões ou estrutura nos dados sem qualquer orientação.

Por último, imagine que você está jogando um vídeo game no qual você obtém pontos ao fazer os movimentos certos. O jogo premia por decisões sábias e pune por decisões erradas, ajudando você a aprender quais ações levam ao sucesso. Isso é parecido com o aprendizado por reforço, no qual a máquina aprende por tentativa e erro, recebendo recompensas por tomar decisões corretas e penalidades

No problema da distinção entre bananas e maçãs, ao utilizar aprendizado supervisionado, as fornecidas fotos para treinamento teriam uma queta (rótulo). As bananas teriam uma etiqueta com a palavra "banana" e as maçãs teriam uma etiqueta com palavra "maçã". Se tirássemos os rótulos, o algoritmo do tipo supervisionado não teria como aprender.



pelas erradas.

Compreender esses tipos de aprendizado irá não apenas aprofundar sua apreciação pela forma como máquinas aprendem, mas também lançar luz sobre a forma como você mesmo aprende e se adapta todos os dias.

2.6.1 Aprendizado supervisionado

O aprendizado supervisionado é um conceito que engloba todos os algoritmos de AM que precisam de dados rotulados para poderem aprender. No aprendizado supervisionado há algoritmos de dois
tipos bem definidos: regressão e classificação. Quando um algoritmo
é um regressor, a previsão feita por ele necessariamente vai se dar na
forma de um valor numérico. Por exemplo, quando queremos determinar o preço de uma casa a partir de suas características utilizaremos
um algoritmo de regressão, já que nosso objetivo é obter (prever), ao
final, um valor numérico (o preço da casa). Um exemplo de algoritmo
de AM regressor é a MVS (Máquina de Vetores de Suporte). Você conhecerá um pouco mais sobre ele no próximo capítulo.

Por sua vez, os algoritmos de AM do tipo classificador realizam previsões categóricas, ou seja, com base nas características do indivíduo eles determinam a qual classe ele pertence. Por exemplo, um algoritmo de AM classificador poderia determinar se o animal presente em uma imagem é um gato ou um coelho com base em certas características; nesse caso, as classes são gato e coelho. Um exemplo de algoritmo de aprendizado de máquina classificador são as árvores de decisão (*Decision Trees*). Aprenderemos mais sobre elas também no próximo capítulo.

Dados rotulados são dados categorizados em classes. exemplo, se você tem um algoritmo de previsão de fraudes em cartão de crédito você utilizaria uma base de dados que contém algumas transações que são fraudulentas e outras que não fraudulentas, você já sabe antecipadamente a qual classe esses dados pertencem.



Podemos associar os algoritmos supervisionados a alunos aprendendo com o auxílio de um professor: a figura do professor está ali para determinar se o estudante está aprendendo ou não, já que os dados fornecidos aos algoritmos supervisionados possuem rótulos, e então os algoritmos sabem exatamente o que eles devem fazer e que tipo de resposta devem fornecer.

2.6.2 Aprendizado não-supervisionado

Os algoritmos de **aprendizado não-supervisionado**, por sua vez, não possuem dados rotulados para poder realizar as previsões; ou seja, pode-se dizer que eles seriam como alunos sem um professor. Você deve estar se perguntando, então, como o algoritmo vai funcionar sem um professor para dizer se ele está acertando ou errando suas previssões. A resposta é simples, mas muito importante porque assinala uma diferença crucial entre o tipo supervisionado, que vimos na seção anterior, e este tipo: A função de um algoritmo não-supervisionado não é realizar previsões como os supervisionados e sim **descobrir padrões nos dados**.

A intenção de algoritmos de aprendizado de máquina não-supervisionados é justamente identificar padrões para conseguir categorizar um conjunto de dados e descobrir quais as semelhanças e as diferenças entre eles e, dessa forma, conseguir dividir um gigante conjunto de dados que aparentam não ter nenhuma correlação entre si em conjuntos de dados menores. O objetivo de um algoritmo desse tipo quase sempre é fornecer uma análise inicial de um conjunto de dados muito grande.

Há três tipos de algoritmos de aprendizado não-supervisionado:



algoritmos de **agrupamento**, algoritmos de **regra de associação**, e algoritmos de **redução de dimensionalidade**.

2.6.3 Aprendizado não-supervisionado: agrupamento

Os algoritmos de **agrupamento** (também conhecido como algoritmos de **clusterização**) têm o objetivo de dividir o conjuntos de dados em subconjuntos. Imagine que você tenha diversas categorias de produtos, mas não sabe a qual categoria cada produto pertence, o algoritmo de agrupamento identificaria os produtos com base em suas características e separaria eles da melhor forma possível.

Esse agrupamento pode ser realizado de forma hierárquica ou de forma particional.

- agrupamento hierárquico: esse tipo de agrupamento assume que os subgrupos contidos nos conjuntos de dados são correlacionados, ou seja, que alguns subgrupos estão dentro de outros ainda maiores, criando assim uma hierarquia entre eles.
- agrupamento particional: esse tipo de agrupamento engloba os algoritmos que têm o objetivo de separar os conjuntos de dados de forma simplificada, ou seja, criando subgrupos independentes.

2.6.4 Aprendizado não-supervisionado: regras de associação

Os **algoritmos de regras de associação** funcionam de uma forma um pouco diferente. Nesse caso, a intenção não é separar os dados e sim entender a forma como eles se relacionam em conjunto. Para ilustrar esse conceito, podemos pensar em um supermercado: se o dono precisa intensificar as vendas de um determinado produto, ele



pode simplesmente descobrir que tipo de produto costuma ser comprado em conjunto com outros. Por exemplo, se é de conhecimento comum que talco normalmente é comprado juntamente com fraldas (pois estão sendo comprados para um bebê), uma forma eficiente de aumentar as vendas de ambos os produtos seria colocando o talco e as fraldas na mesma prateleira ou até mesmo realizar uma promoção de talco se for comprado junto com fraldas. Esse tipo de associação em conjunto é justamente o que as **regras de associação** fazem, para isso elas utilizam entre outras coisas, conceitos de **Probabilidade e Estatística**.

2.6.5 Aprendizado não-supervisionado: redução de dimensionalidade

Normalmente conjuntos de dados não descrevem apenas uma característica de um conjunto de indivíduos. Por exemplo, em um cenário em que você precise realizar a previsão de salários de uma população, você não vai usar apenas a idade de cada pessoa para determinar o salário que ela ganha, pois essa informação não é suficiente. Para determinar o salário, você poderia usar também como critério a profissão, uma vez que certas profissões têm uma faixa salarial maior; ou seja, é um atributo relevante para um modelo mais preciso. Parte da tarefa de um algoritmo de AM é justamente definir quais das características têm um peso maior no resultado final. Esse conceito é conhecido como abstração. Pode parecer simples determinar os atributos mais importantes no exemplo do salário; contudo, se o conjunto de dados tiver muitas variáveis, pode ser uma tarefa complicada.

A tarefa dos algoritmos de redução de dimensionalidade é jus-



tamente reduzir a quantidade de variáveis (dimensões) de um conjunto de dados. Se um conjunto de dados possui cinquenta variáveis, ele terá cinquenta dimensões diferentes, e isso torna o processo de aprendizagem extremamente difícil e pode gerar modelos ruins ou que demorem um tempo muito longo para serem treinados. Esses algoritmos conseguem reduzir as variáveis de tal forma que o aprendizado não seja muito prejudicado.

2.6.6 Aprendizado reforçado

O aprendizado reforçado funciona como se fosse um jogo entre a máquina e o mundo real, no qual a máquina, ou **agente**, aprende a tomar decisões por meio de tentativa e erro a fim de alcançar um determinado objetivo. Imagine que você está aprendendo a jogar um novo vídeo game. No início, você não sabe quais botões pressionar para movimentar seu personagem ou como obter pontos, mas ao longo do tempo você começa a entender o que funciona e o que não funciona por meio da experimentação. Isso é exatamente o que acontece no caso do aprendizado reforçado.

Nesse tipo de aprendizado, o agente interage com o ambiente e recebe recompensas ou penalidades baseadas em suas ações. O objetivo é maximizar as recompensas ao longo do tempo. Por exemplo, em um jogo, a recompensa pode ser dada na forma de pontos ganhos ao se coletar itens ou derrotar inimigos, e a penalidade pode ser perder uma vida ao cair em um precipício. O agente utiliza essas recompensas e penalidades para aprender a melhor estratégia para vencer o jogo.

Uma característica peculiar do aprendizado reforçado é que ele





Figura 4: IA lidando com recompensas e penalidades em um jogo.

não precisa de dados rotulados, tal como ocorre no aprendizado supervisionado. Em vez disso, o agente aprende exclusivamente através da experiência de interação com ambiente. Isso torna essa técnica particularmente útil em situações nas quais é difícil ou impossível fornecer exemplos corretos de antemão, tais como aprender a fazer um robô caminhas ou ensinar um computador a jogar um jogo muito complexo do zero.

Além disso, o aprendizado por reforço é muito flexível e pode ser aplicado em uma variada gama de problemas. Desde jogar xadrez ou "go" até a operação de veículos autônomos e sistemas de gerenciamento de recomendações, o aprendizado por reforço oferece uma maneira muito poderosa para que as máquinas aprendam a tomar decisões inteligentes e adaptar-se a situações novas com pouca ou nenhuma intervenção humana.

Podemos afirmar, de forma resumida, que o aprendizado por reforço consiste em aprender a melhor ação a se tomar em diferentes situações a fim de se atingir um objetivo. Ao longo do tempo, o agente



se torna cada vez melhor em sua tarefa, assim como você consegue se tornar melhor ao jogar vídeo games que você aprende e pratica. É uma forma de AM muito interessante que mostra como as máquinas podem desenvolver habilidades e inteligência de maneira similar àquela adotado pelos seres humanos, usando experiência e prática.

2.7 Separação dos dados

Antes de concluirmos esta apresentação inicial do AM, vamos abordar a divisão dos dados em conjuntos de treino e teste, um conceito fundamental que introduzimos no começo deste capítulo.

2.7.1 Dados de treino

Para que um algoritmo de aprendizado de máquina seja capaz de gerar um resultado satisfatório ele deve passar por uma etapa conhecida como treinamento, que podemos descrever como o processo de ensinar o algoritmo a reconhecer os padrões que desejamos.

Por exemplo, imagine o algoritmo como um aluno que precisa estudar. Nesse caso, a etapa de treinamento seria o estudo. Nessa etapa o algoritmo tem acesso a uma parte dos dados para que possa gerar um modelo capaz de realizar previsões com base no que aprendeu; ou seja, é como se, na intenção de ajudar um aluno a aprender uma matéria, um professor entregasse a ele uma lista de exercícios com as respostas. Dessa forma, o aluno poderia medir o próprio desempenho e determinar onde ele errou com base nos exercícios da lista.



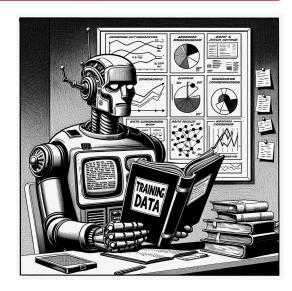


Figura 5: IA aprendendo com dados de treino.

2.7.2 Dados de teste

Por outro lado, temos também os dados teste. Imagine que esses dados sejam como um teste surpresa exigido do aluno após ele ter estudado usando aquelas listas de exerícios. É a oportunidade para que o professor (ou, no caso da IA, nós mesmos) veja se o estudante (no caso da IA, o algoritmo) realmente entendeu a matéria. Os dados de teste são diferentes dos dados de treino porque eles mostram ao algoritmo uma nova informação que ele não viu antes. Da mesma forma, o teste surpresa não vai incluir perguntas iguais à da lista. Isso nos ajuda a descobrir se ele aprendeu corretamente ou se está apenas memorizando as respostas. Então usamos esses novos dados, vemos as respostas que o algoritmo fornece e as comparamos com as respostas corretas. Dessa forma, podemos medir o quão preciso ele está e se necessita de mais treino.



Neste capítulo, mergulhamos nos diferentes tipos de aprendizado que fundamentam o campo do Aprendizado de Máquina. Com essa base estabelecida, estamos prontos para avançar e começar a explorar alguns dos importantes algoritmos mais que fazem a mágica desse tipo de IA acontecer. Está animado(a) para prosseguir nessa aventura de descobertas? Então, siga conosco para o próximo capítulo, onde a jornada do conhecimento continua!

REFERÊNCIAS

- [1] A. Géron. Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow. Alta Books, 2019.
- [2] Kamal Kant Hiran, Ritesh Kumar Jain, Kamlesh Lakhwani, and Ruchi Doshi. Machine Learning: Master Supervised and Unsupervised Learning Algorithms with Real Examples (English Edition). BPB Publications, 2021.
- [3] Batta Mahesh. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)*.[Internet], 9(1):381–386, 2020.
- [4] Clark F Olson. Parallel algorithms for hierarchical clustering. *Parallel computing*, 21(8):1313–1325, 1995.
- [5] FY Osisanwo, JET Akinsola, O Awodele, JO Hinmikaiye, O Olakanmi, J Akinjobi, et al. Supervised machine learning algorithms: classification and comparison. *International Journal of Computer Trends and Technology (IJCTT)*, 48(3):128–138, 2017.