



# LEVANDO IA PARA PRODUÇÃO

## UNIDADE I INTELIGÊNCIA ARTIFICIAL – *MACHINE LEARNING E DEEP LEARNING* E COMO PUBLICAR EM PRODUÇÃO

## **Elaboração**

Paulo Vitor Pereira Cotta

Natasha Sophie Pereira

## **Produção**

Equipe Técnica de Avaliação, Revisão Linguística e Editoração

# SUMÁRIO

## UNIDADE I

INTELIGÊNCIA ARTIFICIAL – <i>MACHINE LEARNING</i> E <i>DEEP LEARNING</i> E COMO PUBLICAR EM PRODUÇÃO .....	5
---	---

## CAPÍTULO 1

INTELIGÊNCIA ARTIFICIAL.....	5
------------------------------	---

REFERÊNCIAS .....	15
-------------------	----



# INTELIGÊNCIA ARTIFICIAL – *MACHINE LEARNING* E *DEEP LEARNING* E COMO PUBLICAR EM PRODUÇÃO

## UNIDADE I

### CAPÍTULO 1

#### INTELIGÊNCIA ARTIFICIAL

##### O que é Inteligência Artificial?

Inteligência Artificial é uma área de estudo que visa à construção de sistemas artificiais que sejam capazes de lidar com problemas de forma automatizada e que possuam capacidade de raciocínio e de tomada de decisões, características antes exercidas apenas pelos seres humanos

Existente há décadas, essa área da ciência é grandemente impulsionada pelo rápido desenvolvimento da informática e da computação, permitindo que novos elementos sejam rapidamente agregados à IA.

Iniciadas nos anos 1940, as pesquisas em torno dessa incipiente ciência eram desenvolvidas apenas para identificar novas funcionalidades para o computador, ainda em projeto. Com o advento da Segunda Guerra Mundial, veio também a necessidade de se impulsionar a indústria bélica por meio do uso de tecnologia.

Com o tempo, surgiram várias linhas de pesquisa dentro do universo da IA, em especial a linha biológica, que visa ao desenvolvimento de conceitos para imitar as redes neurais humanas. Na verdade, nos anos 1960, essa técnica se inicia com a Inteligência Artificial, e os pesquisadores acreditam ser possível que máquinas e algoritmos realizem tarefas humanas complexas como raciocinar.

A IA ficou adormecida por um período, mas, a partir dos anos 2000, voltou a evoluir muito, sendo que a maior evolução e continuidade da IA foi em 2012. Desde então, os algoritmos estão sendo evoluídos, e as empresas estão querendo consumir mais da tecnologia.

Hoje em dia, são várias as aplicações da Inteligência Artificial em nosso cotidiano:

- » jogos;
- » programas de computador;
- » algoritmos de segurança para sistemas;
- » robótica (robôs auxiliares);
- » dispositivos para reconhecimento de escrita à mão e reconhecimento de voz;
- » programas de diagnósticos médicos e outros.

A IA tem uma subárea ou recurso chamado de *Machine Learning* (ML), que nasceu do processo de reconhecer padrões e da teoria de que computadores podem aprender sem serem programados para realizar tarefas específicas. E pesquisadores interessados em IA queriam saber se as máquinas poderiam aprender com dados.

Como é um campo amplo, teremos um capítulo específico para explicar e detalhar as implicações de ML na vida real.

## ***Machine Learning* (ML)**

ML pode ser definida como a prática de usar algoritmos para coletar dados ou extrair *features*, aprender com eles, e, então, fazer uma classificação ou predição sobre alguma coisa no mundo. Então, em vez de implementar as rotinas de *software* manualmente, com um conjunto específico de instruções para completar uma tarefa em particular, a máquina é “treinada” (processo de aprender algo, após a extração das características) usando uma quantidade grande de dados e algoritmos que dão a ela a habilidade de aprender como executar a tarefa.

O processo de treino de um algoritmo de ML é realizado basicamente quando o profissional analisa os dados, verifica e executa a extração de características e, após isso, coloca o modelo para ser treinado.

Para o algoritmo aprender, devem ser aplicados dados supervisionados e ou semisupervisionados. A partir desse processo, é realizada a extração de características, sendo que vários fatores podem ser usados. A seguir, são apresentados os fatores e o detalhamento de cada um:

- » seleção de variáveis: às vezes, o *Dataset* (Conjunto de Dados) é grande e deve ser avaliado por profissionais. Muitas vezes o profissional deve verificar as variáveis

e perceber, de acordo com o seu senso analítico, quais são melhores para treinar o modelo de acordo com o problema que deve ser solucionado;

- » Correlação de Pearson (): também conhecido como correlação linear ou de Pearson, é utilizado para indicar, com valores entre -1 e 1, o grau de relação entre duas variáveis quantitativas. É utilizado para avaliar se uma variável é válida para seu problema;
- » dado supervisionado: é quando o dado já possui um rótulo para os seus dados, tendo a ideia de que existe uma relação entre a entrada e saída;
- » dado semissupervisionado: o aprendizado semissupervisionado é usado onde temos uma grande quantidade de dados de treino, mas apenas alguns são supervisionados.

Transformar os dados de entrada em um conjunto de características, ou *features*, é chamado de extração de características dentro de um conjunto de dados. Se as características extraídas foram cuidadosamente escolhidas, espera-se que esse conjunto represente a parte relevante da informação para se executar a tarefa desejada, em vez de se usar os dados de entrada na íntegra (ALPAYDIN, 2010).

Itens que fazem parte da extração de características:

- » análise de componentes principais em suas características;
- » agrupamento semipreciso (*Semidefinite Embedding*);
- » redução dimensional com múltiplos fatores dentro do conjunto de dados;
- » redução dimensional não linear dentro das camadas;
- » núcleo PCA para retirar as redundâncias;
- » análise semântica latente preparadora para as saídas;
- » análise de componentes independentes dentro das camadas.

O processo com maior custo na criação de uma ML é o de extrair as características para executar no modelo, e, muitas vezes, esse processo pode levar semanas, o que chamamos de fase de preparação dos dados.

Os principais algoritmos de ML são:

- » árvores de decisão: o algoritmo classifica a população para vários conjuntos de dados em algumas propriedades escolhidas usando estatística de primeira ordem de uma população. Esse algoritmo é usado para resolver problemas de classificação;

- » Classificação Naïve Bayes: os classificadores Naïve Bayes são processos de classificação probabilística simples com base na aplicação Bayes, teorema com uma forte independência entre as características. A equação do teorema é , em que:  $P(y)$  é a probabilidade posterior,  $P(x)$  é a probabilidade,  $P(y|x)$  é a probabilidade prévia e  $P(x|y)$  é preditor de probabilidade prévia do Naïve Bayes;
- » regressão linear: é um algoritmo de regressão linear usando os pontos dos dados para encontrar a melhor linha de ajuste. Uma linha é representada pela equação  $y = ax + b$ , em que  $y$  é a variável dependente, e  $x$  é a variável estatisticamente independente. As teorias básicas de cálculo são aplicadas para encontrar os valores para  $a$  e  $b$  usando todas as informações que foram apresentadas no conjunto de dados;
- » SVM (*Support Vector Machine*): pertence ao algoritmo para classificação. O algoritmo executará o processo de separação dos pontos de dados usando uma linha, que é escolhida em uma reta e deve tentar encontrar os dados mais próximos, ou aproximados, usando a distância do menor entre dois pontos;
- » KNN (K-vizinhos próximos): é um algoritmo simples que prevê pontos de dados desconhecidos entre os seus vizinhos mais próximos. O valor de  $K$  é um fator muito crítico quanto à precisão da predição ou classificação. Determina o mais próximo ao calcular a distância entre dois pontos, usando funções básicas de distância como cálculo de distância euclidiana;
- » regressão logística: é utilizada de modo que uma saída discreta é esperada ao identificar uma ocorrência (ex.: prever se a fará sol ou não). No geral, usa alguma função para alocar dados dentro de determinado intervalo de valores;
- » algoritmos de gradiente: *Gradient Boosting Algorithm* usa vários algoritmos com menor poder para criar um algoritmo mais preciso e mais poderoso. Em vez de usar um único estimador, ter vários criará um algoritmo mais estável.

Arthur Samuel descreveu ML como: “campo de estudos que dá aos computadores a capacidade de aprender, sem ser explicitamente programados”, que é uma definição informal mais velha.

Mesmo sendo um campo antigo, muitas coisas novas estão acontecendo no campo de ML. Uma delas é a subárea *Deep Learning* (DL), que trabalha com aprendizados mais profundos e extrai as características de forma automática dentro de seus múltiplos *perceptrons*. No próximo capítulo será abordado o tema de DL.



## Deep Learning (DL)

O *Deep Learning* é uma técnica de *Machine Learning* que utiliza redes neurais profundas para processar informações e aprender com elas. De acordo com Evandro Souza (2018), o aprendizado profundo, do inglês *Deep Learning* (aprendizado estruturado profundo, aprendizado hierárquico ou aprendizado de máquina profundo), tem como base um conjunto de algoritmos que visam modelar abstrações de dados de alto nível usando várias camadas de processamento em uma rede neural, camadas que são compostas por várias transformações lineares e não lineares.

Os algoritmos de DL são representações distribuídas. A suposição subjacente por representações distribuídas é que os dados são observados e são gerados pelas interações de fatores organizados em camadas. A aprendizagem profunda inclui a suposição, predição e classificação, de modo que essas camadas de fatores correspondem a níveis de abstração ou de composição, ou linearmente. Devem ser usadas quantidades e tamanhos de camadas diferentes e para fornecer quantidades diferentes de abstração.

Um dos processos utilizados é o de múltiplos *perceptrons*. Em uma rede neural profunda, utiliza-se o que é conhecido como *backpropagation* (retropropagação), o que ele realmente faz é derivar na ida e na volta até chegar no mínimo do dado. Esse assunto será abordado no próximo tópico juntamente com o *perceptron*. Yann LeCun *et al.* (1989) aplicaram retropropagação supervisionada a essas arquiteturas. Weng *et al.* (1992) publicaram redes neurais convolucionais para o reconhecimento de objetos 3D a partir de imagens de cenas desordenadas e para a segmentação de tais objetos a partir de imagens. Dessa forma, conseguiram estabelecer uma melhora significativa na IA.

É importante entender *Deep Learning* pelo crescimento que essa abordagem tem experimentado nos últimos anos. Ajudou nos avanços em várias áreas, conforme apresentado: identificação dos objetos, a tradução automática de texto e voz, e reconhecimento de todos os tópicos de pesquisa de voz, que por vários anos foram problemas não solucionáveis dentro da academia para pesquisadores de IA.

Em relação à importância do *Deep Learning* ou aprendizagem profunda da máquina, seguem alguns exemplos de como se pode utilizar e de finalidades para as quais ela está sendo constantemente utilizada:

- » desenvolver e analisar uma melhor experiência dos usuários em resultados de pesquisas *on-line*, como UI/UX;
- » otimizar campanhas de *marketing* e anúncios *on-line* em tempo real, *sites* e aplicativos móveis;

- » analisar sentimentos por meio de textos, voz e imagem;
- » melhorar as ofertas ou promoções em e-commerce por meio da análise da navegação do cliente conectado e validação de como o sistema está sendo utilizado;
- » prever falhas, utilizando os logs, em sistemas, em equipamentos diversos e servidores;
- » analisar a identificação e melhorar os preços após a análise do comportamento do consumidor;
- » detectar invasões, problemas ou fraudes, inclusive as invasões de rede orquestradas em massa ou nos servidores;
- » reconhecer padrões, imagens e textos;
- » filtrar *spams* nos *e-mails* e classificá-los;
- » entre outros itens.

#### Principais algoritmos de *Deep Learning*:

- » *GAN (Generative Adversarial Networks)*: o GAN é uma técnica de aprendizado de máquina não supervisionada e de aprendizagem profunda proposta por Ian Goodfellow e alguns outros pesquisadores, incluindo Yoshua Bengio, em 2014;
- » *LSTM (Long short-term memory)*: LSTM é também chamada de “redes de memória de longo prazo”. Relaciona-se com o fato de que os LSTMs são projetados para evitar o problema de dependência de longo prazo, por conta da sua memória de grande volume. Afinal, o objetivo é recordar a informação por longos períodos, o que é praticamente o seu comportamento padrão, nesse caso, algo bem previsível e fácil de aprender;
- » *Autoencoder*: são redes neurais treinadas com a finalidade de transcrever seu *input* para seu *output*. Na prática, o objetivo é aprender as representações dos dados (*encodings*) que podem ser utilizadas na redução da dimensionalidade ou compressão de arquivos.

## **Big Data**

O termo *Big Data* é relativamente novo dentro das empresas, porém o processo de coletar e armazenar grandes quantidades de informações para análises eventuais é antigo e acarreta diversos problemas. O conceito de *Big Data* ganhou força no começo

dos anos 2000, quando o analista Doug Laney (2000) articulou, em 3 Vs, a definição que atualmente é dominante para essa área. São eles:

- I. **Volume:** diz respeito à quantidade de informações coletadas pelas organizações, muitas vezes de fontes variadas, como transações financeiras, comentários em mídias sociais, informações adquiridas por sensores e até mesmo dados que trafegam de máquina para máquina na rede interna. Antes, o armazenamento dessas informações era um problema, mas, atualmente, novas tecnologias auxiliam nesse processo, como é o caso do *Hadoop*.
- II. **Velocidade:** Com o avanço das tecnologias, veio um aumento na velocidade com que os dados são transmitidos, e isso leva à necessidade de tratamento imediato. Essa velocidade pode ser conseguida a partir de etiquetas RFID e até sensores de medições inteligentes.
- III. **Variedade:** além da alta quantidade de dados e da rapidez com que são gerados, ainda existe uma massiva quantidade de diferentes formatos para geração desses dados que os sistemas devem ser capazes de processar. Esses vão desde formatos não estruturados, como é o caso de documentos de texto, *e-mails*, documentos de mídia, transações financeiras, entre outros, até os dados estruturados, que incluem dados numéricos ou armazenados em bancos de dados convencionais.

Doug Laney (2000) apresenta duas características importantes a respeito dos pilares de *Big Data*, que são a **variabilidade** dos dados, visto que, além da crescente velocidade e variedade com que os dados são gerados, seus fluxos podem aumentar de forma inconsistente e com grandes picos periódicos. Qual é a última tendência nas redes sociais? Todos os dias, picos de dados sazonais ou gerados por eventos particulares podem ser difíceis de gerenciar, ainda mais com dados não estruturados; e a **complexidade** dos dados, pois são gerados por múltiplas fontes, dificultando a transição dos dados entre sistemas, ou seja, as ações de limpar os dados, transformá-los, combiná-los, entre outras, tornam-se muito complexas. Para manter o controle sobre os dados, é necessário realizar a conexão e correlação entre os dados, montar hierarquias e múltiplas ligações.

Ainda existem algumas premissas importantes que devem ser levadas em consideração quando se trata de *Big Data*, seus dados e como são utilizados. Por exemplo:

- » **social data:** são dados basicamente oriundos dos usuários e tipos de informações que decifram comportamentos ou padrões pessoais. Nesse caso, é possível identificar perfis para trabalhar de forma mais direcionada para análise de marketing. Quando

há dados de como as pessoas realizam buscas e o que elas comentam nas mídias ou redes sociais, consegue-se perceber o quanto as pessoas são previsíveis;

- » **enterprise data:** são os dados gerados por empresas a todo momento, por exemplo, dados financeiros, de recursos humanos, de operações, entre outros;
- » **personal data ou data of things:** é um novo processo de análise cujos dados são gerados por geladeiras, carros, TVs e outros dispositivos que estão conectados à internet e conversam entre si. É conhecido como Internet das coisas ou *IoT*, tema muito abordado em 2015 e 2016, sendo considerado uma grande tendência e melhoria para os próximos anos. Hoje é possível, por exemplo, pegar as informações do Waze, Google Maps ou outras bases de dados para gerar informações sobre o trânsito em tempo real, o que também é aplicável à saúde, em que é possível coletar informações sobre os batimentos de um paciente e utilizá-las em bases médicas, tudo isso de modo a facilitar a vida humana em seu dia a dia.

A análise de dados é um dos maiores pilares do *Big Data*, ou seja, sem uma análise correta e criteriosa dos dados, é impossível gerar *insights* e direcionar o caminho mais acertado ou correto para o usuário. Por isso, ela é a etapa mais importante do processo em que o Big Data está inserido. O processo da análise passa por inspecionar os dados, avaliar as métricas, verificar as previsões e criar hipóteses para realizar testes com o objetivo de melhorar, verificar ou entender determinado cenário e seus padrões.

O processo de análise dos dados é determinante para encontrar os padrões de comportamento que existem em todo o trabalho de marketing e vendas, o que nos permite monitorar qualquer validação do processo de desvio para algo positivo ou negativo. A análise de dados é responsável por encontrar esses padrões de comportamento para então monitorá-los e, quando houver qualquer mudança, alertar para que seja tomada uma decisão baseada em conhecimentos adquiridos a partir das análises realizadas.

Profissionais com cargos analíticos e estatísticos estão muito requisitados no mercado em diversos setores, como *marketing*, *hardware*, vendas, tecnologia da informação, *e-commerce*, advocacia entre outros.

## Big Data Analytics

De acordo com Rodrigo Nascimento (2017), com o *Big Data Analytics*, extraímos, organizamos, tratamos e compreendemos os dados estruturados e não estruturados. Estamos falando de nada menos que a transformação dos dados em informação útil para pessoas e empresas seguirem em frente com maior segurança, considerando a

necessidade de uma tomada de decisão. Conforme já indicado anteriormente, há 3 tipos de dados – *social data*, *enterprise data* e *personal data* – para explorar em meio às possibilidades que os dados podem nos trazer.

Abaixo, seguem alguns exemplos de onde podemos obter dados para então tratá-los, gerando os *insights* necessários. Quanto às fontes de dados, alguns pontos são importantes para entender o quanto cada uma é empregável para qualquer situação:

- » **e-mails:** o *e-mail* pode ser uma fonte de dados. Alguns sistemas e ferramentas web facilitam e enviam por *e-mail* arquivos com dados de forma padronizada. Essas informações podem vir por anexo via arquivos XLS, PDF ou até mesmo no corpo do *e-mail*. A partir dessas informações, podem ser gerados robôs ou aplicativos que fazem integrações com o *e-mail* a fim de buscar dados e tratá-los de forma a entregar a informação necessária de maneira eficiente e eficaz;
- » **mídias sociais:** mídias sociais como Facebook, Twitter, Instagram, blog e qualquer tipo de redes sociais têm dados valiosos que podem ser extraídos e transformados em informação. Podemos obter o que as pessoas estão falando em mídias sociais sobre assuntos relacionados ao seu mercado ou até mesmo sobre o mercado em si. Os dados das mídias ou redes sociais, como já vimos, são não estruturados e facilitados. Portanto, é necessário realizar organização dos dados e formalizar o mínimo para ser trabalhado;
- » **open data:** cada vez mais os governos estão liberando dados sobre saúde, finanças, consumidores, clima, entre vários outros tópicos e empresas. Podemos usar esses dados para cruzar com fontes de dados internas, avaliar o processo, identificar os pontos e obter informações estratégicas para o negócio, e tomar decisões que antes não eram possíveis por falta de recurso;
- » **sistemas web:** existem diversas ferramentas *on-line* disponíveis para coletar e tratar informações úteis para a empresa, como informações de clientes, empresas, visitantes, oportunidades de negócio, entre outras. Essas informações podem ser coletadas por sistemas web, como é o caso do Google Analytics, RD Station, Facebook Ads, Bing ads e diversos outros. Explorar essas fontes virou opção básica para qualquer empresa.

A importância do *big data* não gira em torno da quantidade de dados que você tem, mas do que você faz com eles. Você pode obter dados de várias fontes e analisá-los para encontrar respostas que permitem 1) reduzir custos; 2) economizar tempo; 3) desenvolver novos produtos e otimizar ofertas; 4) tomar decisões mais inteligentes. (SAS; DOUG LANEY, 2012).

Soluções de *Big Data* são construídas para tratar e trabalhar grandes volumes de dados, sejam estes estruturados ou não estruturados. Isso significa que eles não têm relação entre si nem uma estrutura definida. Seguem alguns exemplos: posts no Facebook, vídeos, fotos, tweets, geolocalização, comportamento, avaliação de mercado, entre outros.

# REFERÊNCIAS

- “Google’s AlphaGo AI wins three-match series against the world’s best Go player”.** TechCrunch. , 25 Maio maio 2017.
- ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKI, T. J. **A learning algorithm for boltzmann machines.** Cognitive science, Elsevier, 1985.
- AIZENBERG, Igor Aizenberg,; AIZENBERG, Naum N. Aizenberg,; JOOS, P.L. **Vandewalle. Multi-Valued and Universal Binary Neurons: Theory, Learning and Applications.** Springer Science & Business Media, 2000.
- BAKER, J.; DENG, Li; GLASS, Jim; KHUDANPUR, S.; LEE, C.-H.; MORGAN, N.; O’SHAUGHNESSY, D. **“Research Developments and Directions in Speech Recognition and Understanding”**, 2009.
- Balázs Csanád Csáji. **Approximation with Artificial Neural Networks.**; Faculty of Sciences; Eötvös Loránd University, Hungary, 2011.
- BENGIO, Y.; COURVILLE, A.; VINCENT, P. **Representation Learning: A Review and New Perspectives.** IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013.
- BENGIO, Yoshua; LECUN, Yann; HINTON, Geoffrey. **“Deep Learning”**. Nature, 2015.
- BOREN, W. L.; HUNTER, T. B.; BJELLAND, J. C.; HUNT, K. R. **Comparison of breast consistency at palpation with breast density at mammography.** Invest Radiol, 1990.
- DENG, L.; YU, D. **“Deep Learning: Methods and Applications” (PDF).** Foundations and Trends in Signal Processing, 2014.
- HINTON, G. E.; SRIVASTAVA, N.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. R. **“Improving neural networks by preventing co-adaptation of feature detectors”**, 2012.
- HINTON, Geoffrey E.; DAYAN, Peter; FREY, Brendan J.; NEAL, Radford. **“The wake-sleep algorithm for unsupervised neural networks”**, 1995.
- HORNIK, Kurt. **“Approximation Capabilities of Multilayer Feedforward Networks”**. Neural Networks, 1991.
- KAPUR, Lenny. **Neural Networks & The Backpropagation Algorithm, Explained.** 1th ed. Spring, 2010.
- KRIZHEVSKY; SUTSKEVER; HINTON. **Methods research image.** 2012.
- LECUN et al., **“Backpropagation Applied to Handwritten Zip Code Recognition,”**, Neural Computation, 1989.
- LU, Z., ; PU, H., ; WANG, F., ; HU, Z., ; & WANG, L. **The Expressive Power of Neural Networks: A View from the Width**, 2017.
- LUI, Bing et al. **Lifelong Machine Learning: Second Edition.** 2018.
- MORGAN, Nelson; BOURLARD, Hervé; RENALS, Steve; COHEN, Michael; FRANCO, Horacio. **“Hybrid neural network/hidden markov model systems for continuous speech recognition”.** International Journal of Pattern Recognition and Artificial Intelligence, 1998.

## REFERÊNCIAS

NASCIMENTO, Rodrigo. **Afinal, o que é Big Data?**. Disponível em: <<http://marketingpordados.com/analise-de-dados/o-que-e-big-data-%F0%9F%A4%96/>>. Publicado em: 2017. Acesso em: 17 de jun. de 2019.

OLSHAUSEN, B. A. *“Emergence of simple-cell receptive field properties by learning a sparse code for natural images”*. Nature, 1996.

SCHMIDHUBER, J. *“Deep Learning in Neural Networks: An Overview”*. Neural Networks, 2015.

SHAVLIK, J. W., and *Dietterich, T. G. Readings in Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1990.

WAIBEL, A.; HANAZAWA, T.; HINTON, G.; SHIKANO, K.; LANG, K. J. *“Phoneme recognition using time-delay neural networks”*, 1989.

WENG, J. WENG, ; AHUJA, N. AHUJA AND; HUANG, T. S. HUANG, *“Learning recognition and segmentation using the Cresceptron”*, 1997.