

CLOUD COMPUTING PARA MACHINE LEARNING: SINERGIA PARA A ESCALABILIDADE E VELOCIDADE DA INTELIGÊNCIA ARTIFICIAL

Welton Henrique Silva Teixeira¹

Julio Cesar Carou Felix de Lima²

Olinda Nogueira Paes Rizzo³

Resumo

Introdução

Cloud Computing e Machine Learning são áreas da computação com grande potencial e atualmente ocupam cada vez mais espaço e que devem se projetar para o futuro. Essas áreas se complementam e apesar de se tratar de campos diferentes, a inteligência artificial precisa trabalhar com grandes quantidades de dados, e com aprendizado em grande escala, com uma quantidade de dados maior ainda. Cloud Computing entrega a escalabilidade necessária com uma infraestrutura dispensando muitas vezes, o aparato físico, trazendo economia, não apenas monetário, mas também, espacial. Existem Large Language Models (LLM) que chegaram absorver mais de 570 gigabytes de apenas dados limpos, ou seja, pós filtragem a partir do Big Data que podia envolver mais de 45 terabytes de conteúdo (SHAIP, 2025). O custo desse armazenamento de forma física, pode ultrapassar 5 mil dólares podendo mais que duplicar o valor, ao passar dos 100 terabytes, estimativa dos modelos mais recentes, contando os gastos com a parte física da computação e a energia elétrica necessária para manter os computadores, esse valor poderia ultrapassar os 300 mil dólares. Com Cloud Computing, esse valor pode ser reduzido drasticamente, já que, manutenção, energia gasta e servidores estão incluídos no serviço. Segundo dados divulgados pelos maiores provedores de nuvem do mundo, a Eficácia de Uso de Energia (Power

Usage Effectiveness - PUE), métrica que relaciona a energia total consumida pelo data center à energia usada exclusivamente pelos equipamentos de TI, entrega valores que pelo menos até 2025 ficam em torno de 1,04 a 2,0. Podem ir de 4% a 50% com base na energia total usada pelos servidores, sendo que, quanto menor a porcentagem, maior eficiência (Barr, 2015). A Aprendizagem de Máquina (Machine Learning) aprimora a inteligência artificial por meio da coleta e análise de dados, buscando fazer a máquina pensar ou agir como humanos, com racionalidade ou imitação. Para isso, são necessários dados de alta qualidade. O aprendizado não se limita a responder perguntas complexas, mas também a gerar previsões e respostas probabilísticas. Um exemplo é quando um usuário pesquisa um produto e, em seguida, recebe anúncios relacionados, refletindo a adaptação da máquina ao seu comportamento. Isso acontece rapidamente por conta da forma que o Machine Learning funciona, ele gera, testa, descarta e refina hipóteses em uma fração de segundos, de forma que nenhum ser humano no mundo seria capaz de fazer. Essas “hipóteses” na verdade são geradas a partir de dados, como visto antes, uma LLM manteve padrões de apenas 1,24% dos dados que lhe foram entregues, e os 98,76% restantes foram úteis para o treinamento, testados e descartados. Mesmo os dados mantidos e incorporados ficam armazenados em servidores de nuvens, tornando o processo muito mais rápido entregando mais produtividade para qualquer que seja o processo (Shaip, 2025). Para entender a capacidade do Cloud Computing e porque é um

¹Graduando em Engenharia da Computação da Universidade Santo Amaro, E-mail: henriquewelton013@gmail.com.

²Professor Mestre, Universidade Santo Amaro, SP. E-mail: jclima@prof.unisa.br.

³Professora Mestra, Universidade Santo Amaro, SP. E-mail: orizzo@prof.unisa.br.

ótimo complemento para Machine Learning, é necessário entender o seu conceito. Cloud Computing basicamente é a abstração de recursos físicos para meios virtuais facilitando também sua utilização, modificando e ajustando hardware, software, plataformas de desenvolvimento e serviço para uma carga de trabalho variável de acordo com as escolhas do cliente, comercializado no formato em que se paga pelo uso (pay-as-you-go). Substituindo então aspectos físicos e ativos da TI por serviços integrados virtualmente. Sendo uma de suas maiores vantagens como suporte ao Machine Learning, o fato de sua elasticidade acompanhar a demanda por meio de um ajuste fino, onde se tem a quantidade necessária, anulando gastos redundantes ou desnecessários.

Objetivos

Objetivo Geral

A presente pesquisa tem como objetivo geral analisar e avaliar a relação entre Cloud Computing e Machine Learning com foco no uso de Cloud Computing para Big Data, visando agilizar e reduzir custos de processos e filtragem de dados.

Objetivos Específicos

- Analisar o funcionamento da elasticidade e escalabilidade de uma nuvem para o armazenamento de grande quantidade de dados.
- Investigar a relação entre Machine Learning e Cloud Computing.
- Analisar a diferença entre Cloud Computing e Virtualização.
- Comparar o desempenho e custo de Cloud Computing com soluções baseadas em virtualização local.
- Realizar experimentos com uma IA construída com Big Data em nuvem.

Metodologia

A pesquisa adota uma abordagem mista, teórica e experimental onde o uso de Cloud Computing será testado junto a Machine Learning para trazer amostras reais de desempenho, com parâmetros reais e comparações. A busca pelas informações foi feita a partir de questões que levaram a temas pertinentes, cada uma dessas questões levava a uma ou mais informações, essa informação poderia concluir o assunto ou gerar uma nova questão, assim podendo ramificar-se o tema, semelhante ao funcionamento de uma árvore binária. Diante de informações sobre certas tecnologias, foi buscado uso real de empresas como a NVIDIA, a Pure Storage, a Dell EMC e algumas outras. Houve também a busca de documentações de desenvolvimento com Machine Learning e de serviços de nuvem como a documentação Python e os serviços da AWS. A pesquisa inicial, que descreve informações e procura questões sobre todos os temas que se conectam ao tema principal foi feita a partir de leituras de trabalhos de autores da área, artigos científicos de grandes instituições de ensino e a partir de documentações de grandes corporações de tecnologia. A AWS (Amazon Web Services), plataforma de vendas de serviços de Cloud Computing, forneceu muitas das métricas e informações específicas do funcionamento da Machine Learning dentro do Cloud Computing. O Google Cloud (serviço de nuvem da Google), foi o serviço usado para o experimento. Para conceitos ou dados gerais, documentações e artigos de corporações de tecnologia como Cloudflare, Oracle e NVIDIA, foram utilizados.

Modelo Experimental para Avaliação da Escalabilidade e Elasticidade

O que gerou os conceitos básicos do experimento foi o livro “Data Science do Zero – Noções Fundamentais com Python”, escrito por Joel Grus (engenheiro de Software, cientista de dados e engenheiro pesquisador). Joel Grus propõe um desafio relacionado a ciências de dados. No desafio devemos procurar conectores-chave a partir dos dados de funcionários de uma empresa fictícia, buscando

suas amizades e interesses. O desafio é prático e de certa complexidade, mas define muito bem um conceito de machine learning, o “sistema de recomendação”. Precisou ser construída uma Interface de Programação de Aplicações (API), para que fosse possível criar uma comunicação entre os dados de nomes de funcionários, seus relacionamentos e seus interesses.

Para escalar o processo e testarmos e compararmos a elasticidade da nuvem, foram aumentados o número de funcionários e de interesses, logo após, foi feito um segundo experimento com um sistema de adição de funcionários. A Inteligência artificial deveria coletar os dados e modificar as recomendações em tempo real. Todo processo, foi construído utilizando-se do Anaconda (distribuição do Python que inclui o gerenciador de pacotes, incluindo pacotes de Machine Learning), Pycharm (ambiente de Desenvolvimento Integrado para programação em Python), o Google Cloud (serviço de nuvem da Google) e um computador modesto com as seguintes especificações. Processador: Ryzen 7 – 5700g, 16GB de RAM, SSD NVMe PCI Express 3.0.

Após o experimento, precisamos entender se o que ocorreu, também acontecia em maior escala, para isso foram encontrados dados públicos, divulgados por Jennifer Han (engenheira Sênior de Análise na Netflix) e Pallavi Phadnis (engenheira de dados na Netflix), e dados divulgados pelo Google sobre o Spotify.

Resultados e Discussão

Foi feito um experimento simples replicando e adaptando uma situação descrita por Joel Grus, em seu livro, “Data Science do Zero – Noções Fundamentais com Python”, tal experimento se tratava em definir uma lista hipotética de pessoas, pessoas que trabalhem na mesma empresa, após isso, cada uma dessas pessoas definiam suas relações com as outras pessoas da mesma empresa, indo grau um ao dois, grau um significando pouca ou nenhuma proximidade, o grau dois significando proximidade ou amizade entre os funcionários, gerando relações de um para muitos ou de um para um, além disso cada funcionário deveria ter seus próprios interesses como música, arte,

academia ou natureza, gerando dados como os descritos no Quadro 1.

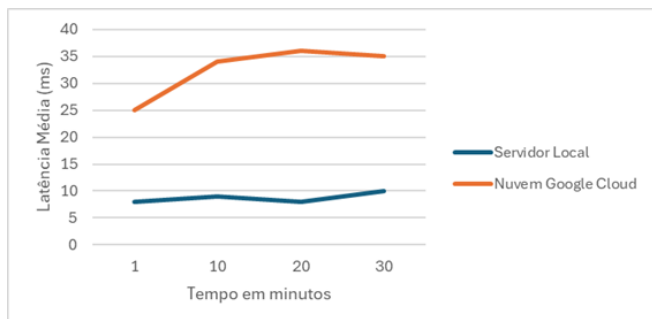
Quadro 1 – Quadro hipotético de funcionários, relacionamentos e interesses

FUNCIONÁRIOS	RELAÇÕES DE PROXIMIDADE	INTERESSES
Alberto	Luciana, João, Hugo	Música, carros, academia
Alexandre	Augusto, Daniela, João	Video-game, academia, natureza
Augusto	Alexandre, Carlos	Carros, esportes
Carlos	Augusto, Daniela	Ufologia, video-game, natureza
Daniela	Augusto, Carlos	Yoga, natureza, academia
Hugo	Alberto, Rosana, João, Pedro	Academia, esportes, natureza
João	Alberto, Alexandre, Hugo	Video-game, ufologia, música
Luciana	Alberto, Pedro	Artes, viagem, natureza
Pedro	Hugo, Luciana	Carros, música, video-game
Rosana	Hugo	Música, esportes, yoga

Fonte: Autoria própria.

Com os dados apresentados no Quadro 1 é possível testar e evidenciar vários dos assuntos já discutidos, apesar de ser um algoritmo simples, ele nos traz a fácil escalabilidade para testar a elasticidade da nuvem, aumentando o número de funcionários, de relações e de interesses. Possibilitando também o descarte de alguns dados menos relevantes, como “com quais funcionários, cada funcionário não tem nenhum tipo de relação?”. Esse sistema criado é uma simplificação de como realmente funciona um sistema de recomendação de uma rede social, ou de um aplicativo de compras. Vemos no quadro de exemplos que Hugo é a única conexão próxima de Rosana, em contrapartida, Hugo também é próximo de Alberto, João e Pedro. A inteligência artificial de uma rede social, traria essas informações da nuvem e passaria a recomendar os amigos de Hugo para Rosana. Os interesses têm função parecida, fariam com que além dos amigos de Hugo fossem recomendados a ela, os funcionários em geral que tivessem interesses parecidos também fossem recomendados, trazendo exemplos reais de filtragem colaborativa e filtragem baseada em conteúdo. Para testar e comparar a elasticidade da nuvem, utilizamos o mesmo método utilizando funcionários, relações e interesses. A Inteligência artificial deveria coletar os dados e modificar as recomendações em tempo real. Após isso, foi testado a diferença no tempo de resposta da IA ao aumentar o número de dados absorvidos em uma nuvem do Google Cloud. E em um servidor local emulado em um computador. Os dados sem a adição dos funcionários a cada minuto geraram a Figura 1:

Figura 1 – Gráfico de comparação de desempenho entre um Servidor local e uma nuvem.

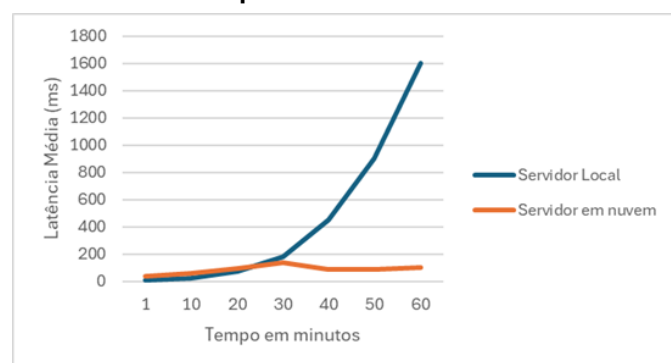


Fonte: Autoria própria.

Esse gráfico demonstra a relação de tempo e latência do funcionamento da IA com os mesmos 100 funcionários pré-estabelecidos, é possível perceber que a latência do servidor local emulado no próprio computador é mais estável e menor que a latência da nuvem. Isso porque, para dados são estáticos e estão em menor quantidade, junto com as oscilações de rede, causadas por problemas de rede.

Ao adicionar 5 funcionários por minuto em um período de uma hora, totalizamos 400 funcionários e a situação se inverte, como pode ser visto na Figura 2, logo abaixo:

Figura 2 - Gráfico de comparação de desempenho entre um Servidor local e uma nuvem após a adição de 5 funcionários por minutos.



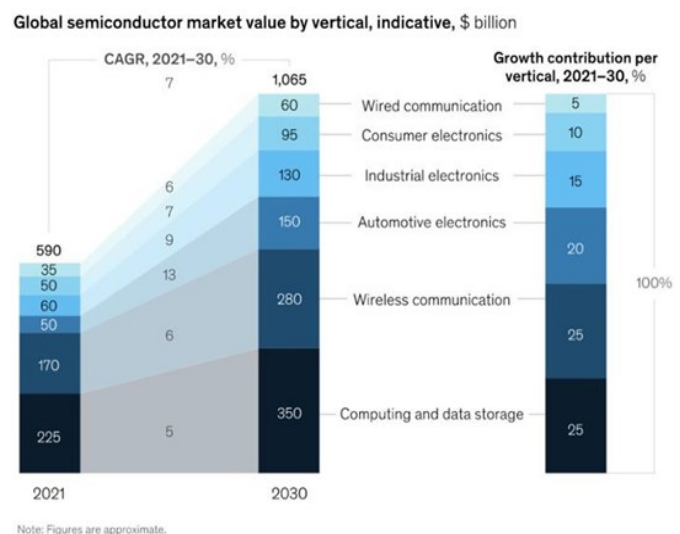
Fonte: Autoria Própria.

Em um tempo menor que 30 minutos é possível perceber que o servidor local mantém uma latência menor, mas logo depois que o número de funcionários passou de 250, a latência chegou até 1600ms (equivalente a 1.6 segundos de espera) no servidor local. Enquanto isso a latência média da nuvem se manteve

parecida durante todo o processo. Isso acontece pois, após 40 minutos, a nuvem ativou o sistema de escalonamento automático, que aloca recursos computacionais de acordo com a demanda do processo. A análise dos dados quantitativos evidencia que para processos dinâmicos, onde a quantidade de dados tende a aumentar exponencialmente a escalabilidade da nuvem tende a se comportar melhor em todo o sistema, entregando mais velocidade de resposta. O funcionamento do servidor local, ao ser colocado sobre estresse, teve uma piora de até 160 vezes comparado ao seu desempenho anterior, com dados estáticos, enquanto a nuvem manteve desempenho parecido e quando se mostrou necessário, utilizou-se de sua elasticidade para manter o desempenho. Isso acontece porque a nuvem utiliza um método chamado de elasticidade horizontal, onde a necessidade responde ao uso e ao preço com o método “paguepelo-uso”. Dividindo a carga de trabalho entre os vários servidores usados pela nuvem. Refinando o processo, utilizando imagens como dados, a diferença seria ainda maior comparado ao computador utilizado como servidor local, isso porque como dito em suas especificações, é utilizado uma placa gráfica integrada, o que pode diminuir ainda mais, o desempenho do servidor local. A melhor forma de manter o desempenho da nuvem e do servidor local mais parelhos, seria, utilizando-se da elasticidade vertical no servidor local. Ou seja, melhorar o servidor emulado no computador local a partir da melhora do hardware do próprio computador. Para comparar essa opção, é preciso buscar a contribuição de cada uma das peças para com Machine Learning e se continua sendo uma opção viável em comparação a nuvem. O processador (CPU), utiliza da sua quantidade de núcleos para executar os processos, como no caso em questão em que o processador contém 8 núcleos para executar cálculos e operações complexas de forma separada, e cada um deles executando uma sequência por vez. O problema é que para tarefas que exigem muito mais como o processamento de imagens ou de modelos de três dimensões, o CPU necessita do acompanhamento de uma GPU. Agindo em conjunto, a CPU busca concluir as tarefas da forma mais rápida possível,

utilizando-se inclusive de memória cache para recuperar dados para a memória principal, agindo sempre em baixa latência. A GPU é construída para suportar a alta latência das memórias e isso a torna essencial para trabalhar em servidores locais, quando os dados são mais que conjuntos de texto. Essas duas peças evidenciam o segundo problema da adoção de servidores locais, junto a peças de armazenamento, memórias RAMs e fontes. O processo é muito mais caro do que contratar o serviço de uma nuvem. No caso das GPUs e CPUs em específico o problema é ainda maior, parte do encarecimento, se deve a projeção de chipsets e sua construção. Existe uma escassez global de semicondutores causada por uma grande demanda por circuitos integrados e pela guerra comercial entre China e Estados Unidos. Segundo Rizzo (2024), a cadeia de fornecimento de silício evoluiu para ser eficiente, mas isso também a tornou frágil. A McKinsey, empresa de consultoria empresarial americana, realizou uma pesquisa mostrando a projeção de valor de mercado para os semicondutores em diversas áreas, de 2021 a 2030. Como pode ser visto na Figura 3.

Figura 3 - A projeção de valor de mercado para os



semicondutores em diversas áreas, de 2021 a 2030.

Fonte: McKinsey, 2024

Apesar de apresentar crescimento, a pesquisa também projetou um aumento de preço dos

chipsets de pelo menos 2% ao ano, o que torna ainda maior a precificação de CPUs e GPUs que depende diretamente de chips para seu funcionamento. Podemos também usar o preço das peças do computador usado para criar o servidor emulado, além de somar o preço de uma elasticidade vertical feita para comparar diretamente com as médias de latência da nuvem. O que nos leva a Tabela 1, com valores médios do mês de setembro de 2025 para peças de computadores e para serviços de nuvem, no padrão de moeda Real Brasileiro.

Tabela 1 – Comparação de gastos entre hardware e nuvem para o experimento (baseado na média de valores de setembro de 2025).

Computador	Computador pós elasticidade	Nuvem valor mensal
R\$ 3.000,00	R\$ 8.000,00	R\$ 200,00

Fonte: Autoria própria.

É preciso um investimento de 5 mil reais para atingir desempenho parecido ao de uma nuvem de 200 reais que oferece suporte e automatização de alguns processos. Para um projeto real, os valores costumam passar dezenas de milhares mesmo para nuvens, clusters (aglomerado de máquinas conectadas em rede que executam serviços de rede) podem chegar a custar dezenas de milhares de dólares (Bogush, 2024).

Servidores locais, além de carregarem o problema de precificação, consomem mais energia elétrica, dados divulgados pela AWS em 2015, utilizando a métrica Eficácia de Uso de Energia (PUE) dizem que clientes de nuvem consomem 84% menos energia comparado a data centers e servidores locais. Enquanto a média do PUE de um data-center local chega a 1,7. A média de provedores de nuvem (contando mesmo os menos eficientes) é de 1.2. Cálculos simplificados nos levaria a uma média de 29% e uma diferença de apenas 0,493 entre os dois valores. Mas o experimento de diferença de custo contra a diferença de eficiência feito anteriormente, foi reforçado pelos dados da AWS. A nuvem requer apenas 23% do número necessário de servidores para a mesma carga de trabalho, e seus servidores são 71% mais

eficientes, chegando ao valor de 16% de gastos de energia em comparação a infraestrutura local. Empresas como a Netflix já migraram todos os centros de dados para nuvem, incluindo aqueles que alimentam as IAs que formam o feed de um usuário com filmes e séries semelhantes a seus gostos e populares em sua região, utilizando os próprios serviços da AWS. Jennifer Han e Pallavi Phadnis, juntos a sua equipe de Engenharia de Ciências de Dados desenvolveram uma solução de dois componentes que atendem corretamente às suas necessidades analíticas, de forma que tudo se mantenha escalável. O primeiro desses componentes é chamado de Dados de Plataforma Fundamentais (Foundational Platform Data - FPD), responsável por fornecer uma camada de dados centralizada para cada os dados da plataforma, organizando e padronizando, criando assim um modelo de dados consistentes. O segundo componente é chamado de Análise de Eficiência em Nuvem (Cloud Efficiency Analytics - CEA), ele foi desenvolvido a partir do FPD, e ele entrega uma camada de dados analíticos com métricas de eficiência da nuvem (Han; Phadnis, 2024). Como evidenciado nos diversos exemplos entregues aqui, a nuvem consegue resolver e tratar dos três grandes desafios de Big Data, volume, velocidade e variedade de dados (Ávila; Souza; Gonzales, 2019). Machine Learning evidencia que os três pontos devem estar correlacionados e dependentes, independentemente do volume de dados é necessário que exista variedade nos dados obtidos. A nuvem garante que mesmo que o volume de dados seja alto, a velocidade também continuará estável. A Netflix não foi a única a reconstruir sua infraestrutura de dados em Cloud, foi possível perceber essa mudança em companhias como Spotify, que com o grande crescimento, precisou recorrer a novos métodos para diminuição de custos, parando de comprar e licenciar data-centers para utilizar do sistema de nuvem entregue pela Google Cloud. Tanto o Spotify quanto a Netflix utilizam sistemas de Machine Learning como parte essencial de seus modelos de negócio. No Spotify, a tecnologia é aplicada em recursos como a “Descobertas da

Semana”, que recomenda músicas com base no histórico de audições e em perfis de usuários com gostos semelhantes. Já na Netflix, o Machine Learning é responsável por gerar recomendações personalizadas de conteúdo, ajustar automaticamente a qualidade dos vídeos e adaptar a interface de acordo com as preferências individuais dos usuários. Ambas têm estruturas parecidas com a do experimento demonstrado anteriormente no artigo. Feitas de uma forma mega escalável, e apesar da diferença de imensidão entre o experimento e as duas aplicações, todas tiveram problemas de custo elevado em relação aos servidores locais. Segundo a AWS (2024) serviços de nuvem robustos também entregam um conjunto de práticas feitas especialmente para facilitar a construção de projetos de Machine Learning, chamado de Operações de Machine Learning (MLOps), construindo um ciclo de vida duradouro com processos interligados, para Machine Learning. Em contrapartida, a adoção do método, não é recomendada para projetos menores.

Considerações Finais

Machine Learning e Cloud Computing são ferramentas complementares, o uso de Cloud Computing para Machine Learning se torna essencial a partir do ponto em que se mostra necessário construir um negócio escalável gastando quantias menores de recursos monetários. Servidores têm evoluído bastante com tecnologias de processamento rápido, inovações de resfriamento, uma escalabilidade melhor do que a de antes, ainda não é possível competir em uma relação “custo x benefício” para trabalhar com Machine Learning, contra a Cloud Computing. Empresas gigantes como Amazon, Google, Netflix e Spotify, crescem exponencialmente todos os dias, o número de dados recebidos, trabalhados e enviados por cada uma delas crescem de forma maior ainda, para evitar que os custos escalassem junto, elas mesmos optaram por mudar parte, ou 100% dos seus ecossistemas para nuvem. Apesar da defesa em relação ao uso de Cloud Computing para Machine Learning,

ning, os servidores possuem grandes vantagens em outras áreas que fazem seu uso, ele entrega o controle absoluto de sua construção e funcionamento, além de envios mais rápidos quando se trata de envios de dados por rede local (Local Access Network - LAN). Por isso, é muito comum encontrar o uso de servidores em sistemas de custos previsíveis (onde os dados não escalonam), ou em casos em que se é trabalhado com dados extremamente sensíveis, acompanhado a regulamentos de restrição de rede e de segurança. Em muitos projetos adota-se a nuvem híbrida, que combina escalabilidade e elasticidade da nuvem com servidores locais para dados sensíveis, ajudando em situações como picos de demanda. Ainda assim, a maioria dos projetos de Machine Learning utiliza Cloud Computing, pois oferece alto poder computacional, lida com Big Data e viabiliza soluções de MLOps, consolidando sua importância nesse contexto.

Palavras-chave

Big Data; Servidores; Serviços; Elasticidade; Nuvem.

Referências

ÁVILA, Daniel Martínez; SOUZA, Edna Alves de; GONZALES, Maria Eunice Quilici. Big Data: continuidade ou evolução. São Paulo: Cultura Acadêmica Editora, 2019.

AWS. AWS Decision Guide: choosing an aws machine learning service. Choosing an AWS machine learning service. 2024. Disponível em: <https://docs.aws.amazon.com/decision-guides/latest/machine-learning-on-aws-howto-choose>. Acesso em: 20 ago. 2025.

BARR, Jeff. Cloud Computing, Server Utilization, & the Environment. Aws Blog. (Online), 05 jun. 2015. Disponível em: <https://aws.amazon.com/pt/blogs/aws/cloudcomputing-server-utilization-the-environment/>. Acesso em: 26 set. 2025.

BOGUSCH, Kevin. Custos de computação em nuvem em 2024. Oracle. (Online), p. 0-0. 24 jan. 2024. Disponível em: <https://www.oracle.com/br/cloud/cloud-computingcost/>. Acesso em: 26 set. 2025.

GRUS, Joel. Data Science do Zero: noções fundamentais com python. 2. ed. Rio de Janeiro: Alta Books Editora, 2020. cap. 1, p. 1-15.

HAN, Jennifer; PHADNIS, Pallavi. Cloud Efficiency at Netflix. Netflix Technology Blog. (Online), 17 dez. 2024. Disponível em: <https://netflixtechblog.com/cloudefficiency-at-netflix-f2a142955f83>. Acesso em: 25 set. 2025.

JACKSON, Fiona. Global Chip Shortage: Everything You Need to Know. Techrepublic. (Online), 6 set. 2024. Disponível em: <https://www.techrepublic.com/article/global-chip-shortage-cheat-sheet/>. Acesso em: 25 set. 2025.

ORACLE. O que é Computação de Alto Desempenho (HPC)? 2022. Disponível em: <https://www.oracle.com/br/cloud/hpc/what-is-hpc/>. Acesso em: 20 ago. 2025

SARREL, Matt. The Semiconductor Crisis: Addressing Chip Shortages And Security. Aerospike. (Online), 17 ago. 2024. Disponível em: <https://aerospike.com/blog/cpu-vs-gpu/>. Acesso em: 25 set. 2025.

SHAIP. Large Language Models (LLM): Guia Completo em 2025: tudo o que você precisa saber sobre Llm. Tudo o que você precisa saber sobre LLM. 2025. Disponível em: <https://pt.shaip.com/blog/a-guide-large-language-model-llm/>. Acesso em: 20 ago. 2025.

WISHART-SMITH, Heather. The Semiconductor Crisis: Addressing Chip Shortages And Security. Forbes. (Online), 19 jun. 2024. Disponível em: <https://www.forbes.com/sites/heatherwishartsmith/2024/07/19/the-semiconductorcrisis-addressing-chip-shortages-and-security>. Acesso em: 25 set. 2025.