



Estatística e Algoritmos na Educação:

yes, we can!

Laboratório de Computação Reconfigurável (LCR)
Instituto de Ciências Matemática e de Computação (ICMC)
Universidade de São Paulo (USP)
gesielrios@usp.br

Hello!

Gesiel Rios Lopes

- **Licenciado em:**
 - Matemática pela UEMA
 - Computação pela UESPI
- **Especialista em:**
 - Matemática e Estatística pela UFLA
 - Análise e Desenvolvimento de Sistemas pela UNIFOR
- **Mestre em Informática Aplicada pela UNIFOR**
- **Doutorando do Instituto de Computação e Matemática Computacional - ICMC**





Uma pequena introdução a era digital

Introdução

- © Vivemos em uma época digital onde existe uma enorme quantidade de dados sendo produzidas, onde a quantidade de informações produzidas atualmente é superior a quantidade de informação que uma pessoa conseguiria lidar durante toda sua vida (Marquesone, 2016).
- © Este cenário proporciona um mundo cheio de desafios importantes no armazenamento, manipulação e análise de forma inteligente (Goldman et al., 2012).

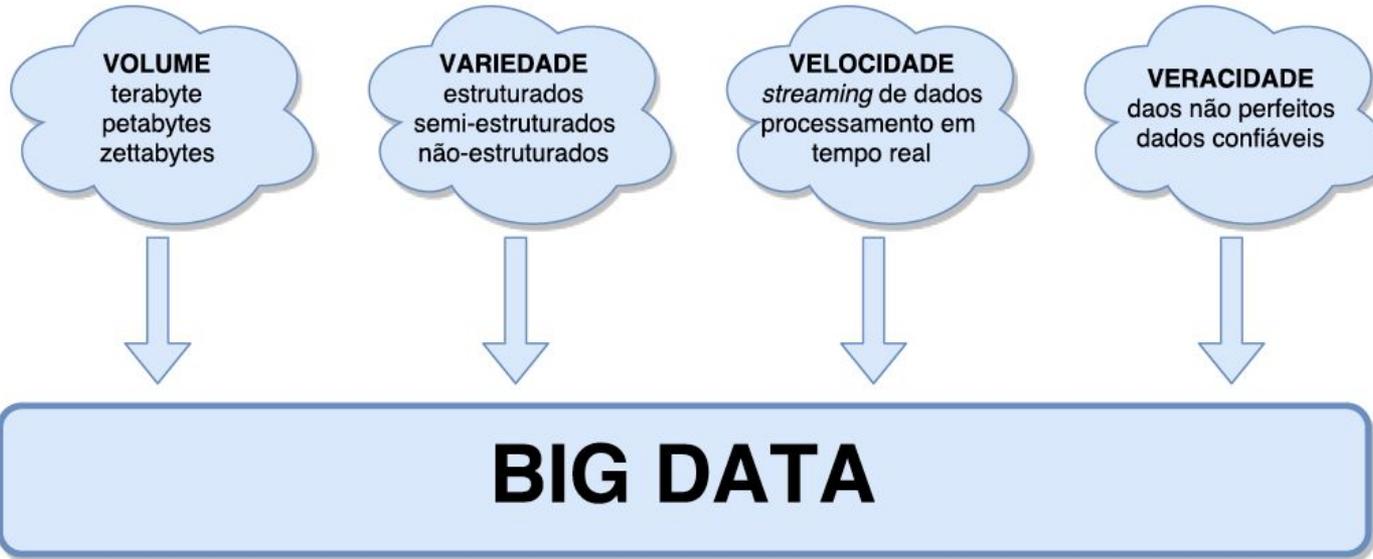
Introdução

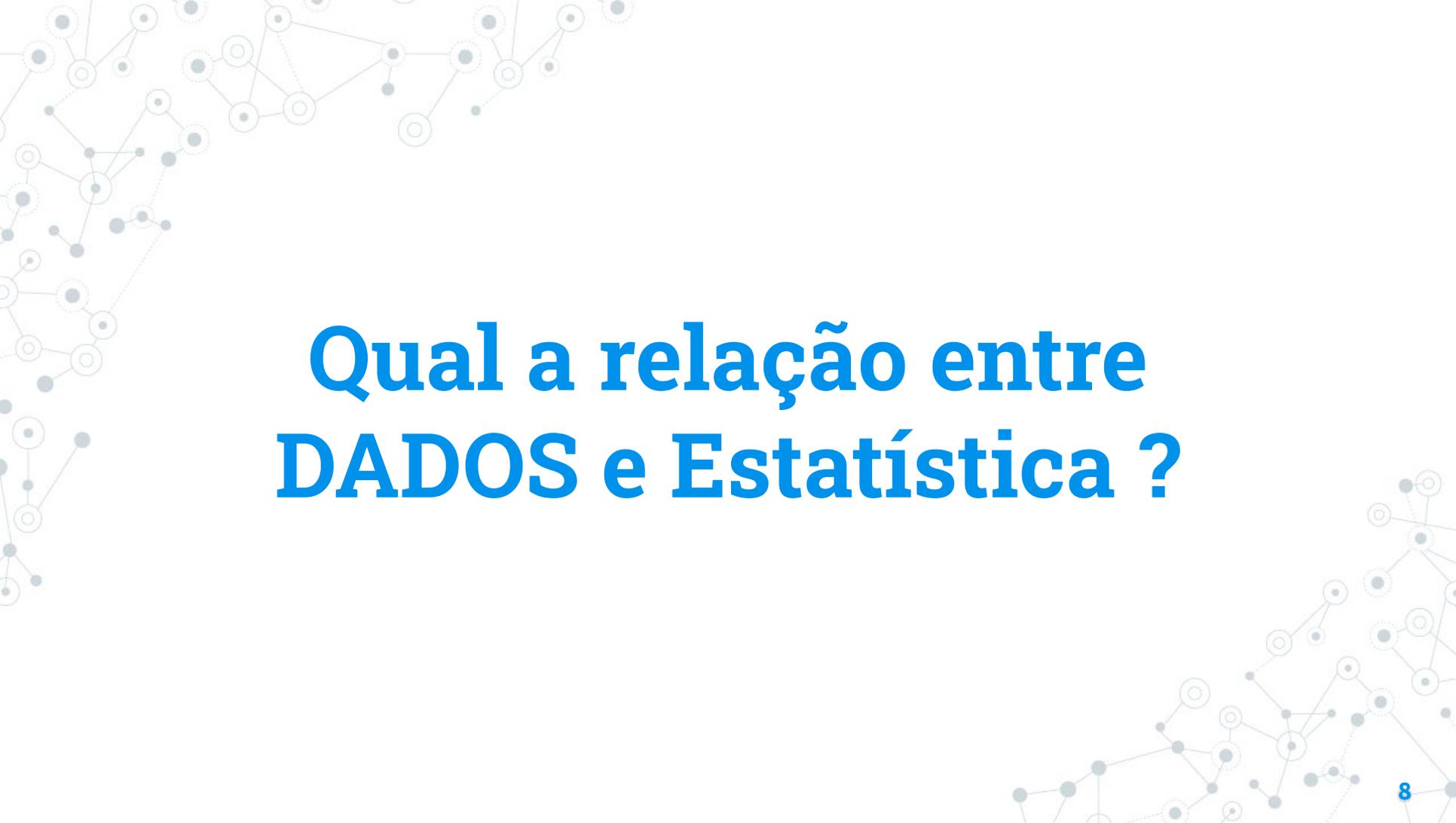
- ◎ Baseado nessas e em outras aplicações que possuem um volume gigantesco de dados, surgiu o conceito denominado “Big Data”.
- ◎ Esse termo faz menção não apenas ao volume, mas também à sua variedade e a velocidade necessária para o seu processamento.

Introdução

- ◎ Big Data é um termo usado principalmente para descrever os conjuntos de dados que são muito grandes e complexos e que requerem de tecnologias avançadas de armazenamento, gestão, análise e visualização (Marquesone, 2016).

Introdução





Qual a relação entre **DADOS** e Estatística ?

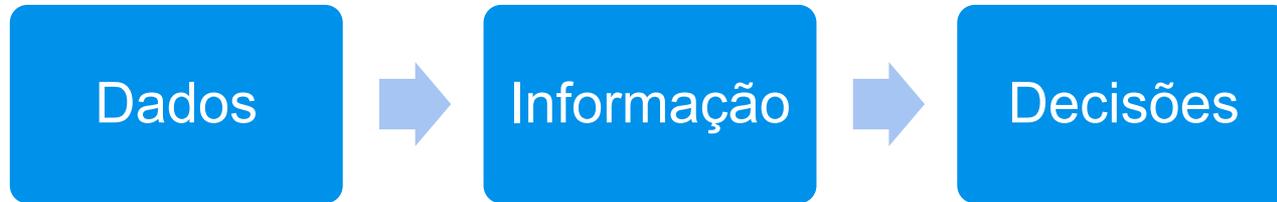
O que é ESTATÍSTICA

- ⊙ “Estatística é a Ciência de obter conclusões a partir de dados”.

Paul Velleman

O que é ESTATÍSTICA

- ◎ A *Estatística* envolve técnicas para coletar, organizar, descrever, analisar e interpretar *dados*, ou provenientes de experimentos, ou vindos de estudos observacionais.



O que é ESTATÍSTICA

⊙ Parte de perguntas/desafios do mundo REAL:

- cientistas querem verificar se uma nova vacina contra febre amarela faz efeito.
- um político quer saber qual é o percentual de eleitores que pretende votar nele nas próximas eleições.
- a Ford quer verificar a qualidade de um lote inteiro de peças fornecidas através de uma pequena amostra.

⊙ Mudança de enfoque:

Atividade de estado para aplicação em pesquisa científica em geral.

Eras da ESTATÍSTICA

◎ A história da Estatística pode ser dividida em três eras:

1. A era de Quetelet (astrônomo, matemático, estatístico belga, 1796-1874) e seus sucessores, cujo o objetivo era obter grandes conjuntos de dados(censos) em ciências sociais.

Eras da ESTATÍSTICA

- ◎ A história da Estatística pode ser dividida em três eras:
 3. A era da produção de dados em massa (*big data sets*), com novas tecnologias dados de alta frequência em finanças, dados astronômicos etc.

Estatística e Computação

- ◎ Avanços em Estatística diretamente relacionados com avanços na área computacional.
 - → 1960: máquinas de calcular manuais, elétricas, eletrônicas.
 - 1960 → 1980: “grandes computadores”
 - 1980 → : computadores pessoais; supercomputadores; computação paralela; “clouds”;
- Era do “*Big Data*” e da “*Data Science*” .

Estatística, algoritmos e inferência

- ◎ Conjuntos de dados enormes (*Big Data*) requerem novas metodologias.
- ◎ Esta demanda está sendo atendida por algoritmos estatísticos baseados em computação intensiva.

Algoritmos

- O que os estatísticos fazem

Inferência

- Porque os estatísticos usam os algoritmos

A decorative network diagram in the top-left corner, consisting of various sized grey circles (nodes) connected by thin grey lines (edges). Some nodes are solid grey, while others are hollow circles with a grey outline. The network is dense and irregular, extending from the top-left towards the center of the slide.

O que é Data Science ?

Data Science

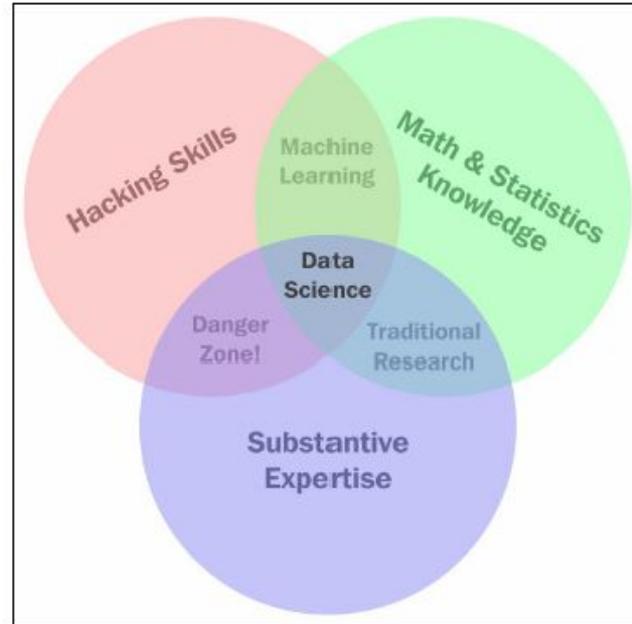
- ◎ Data Science é a arte e a ciência de adquirir conhecimento por meio de dados.
- ◎ Data Science trata de como pegamos dados, os usamos para adquirir conhecimento e, em seguida, usamos esse conhecimento para fazer o seguinte:
 - Tomar decisões
 - Prever o futuro
 - Compreender o passado/presente
 - Criar novas indústrias/produtos

Data Science



- ◎ O entendimento da Data Science começa com três áreas básicas:
 - Matemática/estatística
 - Programação de computador
 - Conhecimento de domínio
- 

Data Science



O diagrama de Venn da Data Science (OZDEMIR, 2018)

Data Science

- ◎ O termo Data Science não é uma terminologia nova:
 - Na década de 1980 C.F. Jeff Wu já sugeria que se adotassem os rótulos *Statistical Data Science*, ou simplesmente, *Data Science*, em lugar de *Statistics*, para dar maior visibilidade ao trabalho dos estatísticos.
 - Tukey (1962, 1977), sob a denominação *Exploratory Data Analysis*, foi o primeiro a dar importância ao que hoje se chama Data Science, sugerindo que se desse mais ênfase ao uso de tabelas, gráficos e outros dispositivos para uma análise preliminar de dados.

Data Science: Perspectiva da Estatística

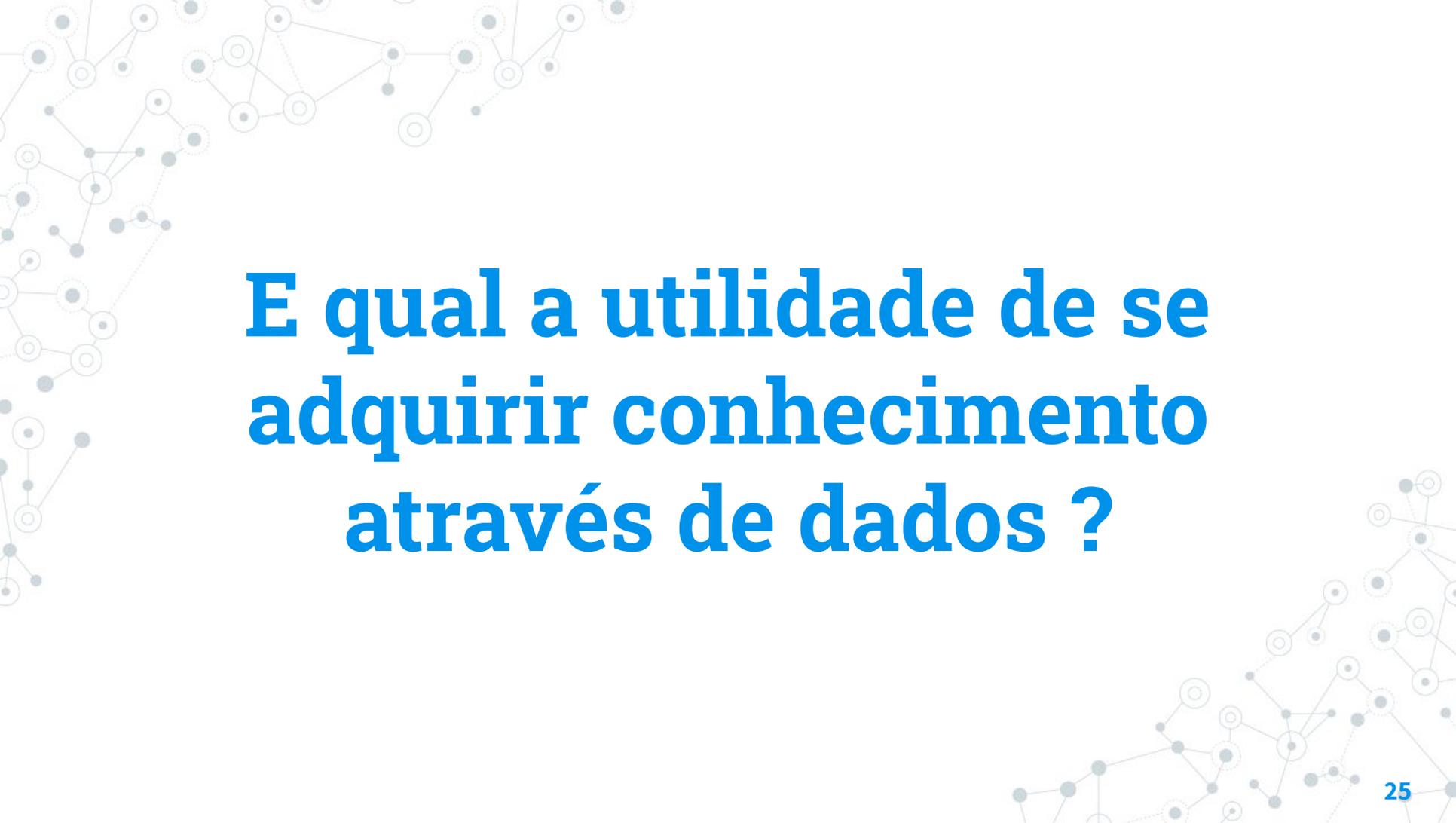
- ⊙ Estatística “serve” a Ciência guiando na coleta e análise de dados.
- ⊙ Dados envolvem incertezas.
 - como foram coletados, medidos ou como foram gerados.
- ⊙ A modelagem estatística ajuda a quantificar e racionalizar incertezas de maneira sistemática.

Data Science: Perspectiva da Estatística

- ◎ Conjuntos de dados são complexos: tipos diferentes de dependência (ao longo do tempo, sobre escalas espaciais, entre variáveis diferentes)
- ◎ Dados de alta dimensão: medimos milhares de variáveis para cada unidade amostral.

Data Science: Perspectiva da Estatística

- ◎ Data Science liga modelos estatísticos e métodos computacionais para resolver problemas específicos de outras disciplinas.
- ◎ Entender o domínio de um problema, decidir quais dados obter, como processá-los, explorar e visualizar os dados, selecionar um modelo estatístico e métodos computacionais apropriados, comunicar os resultados da análise.



E qual a utilidade de se adquirir conhecimento através de dados ?

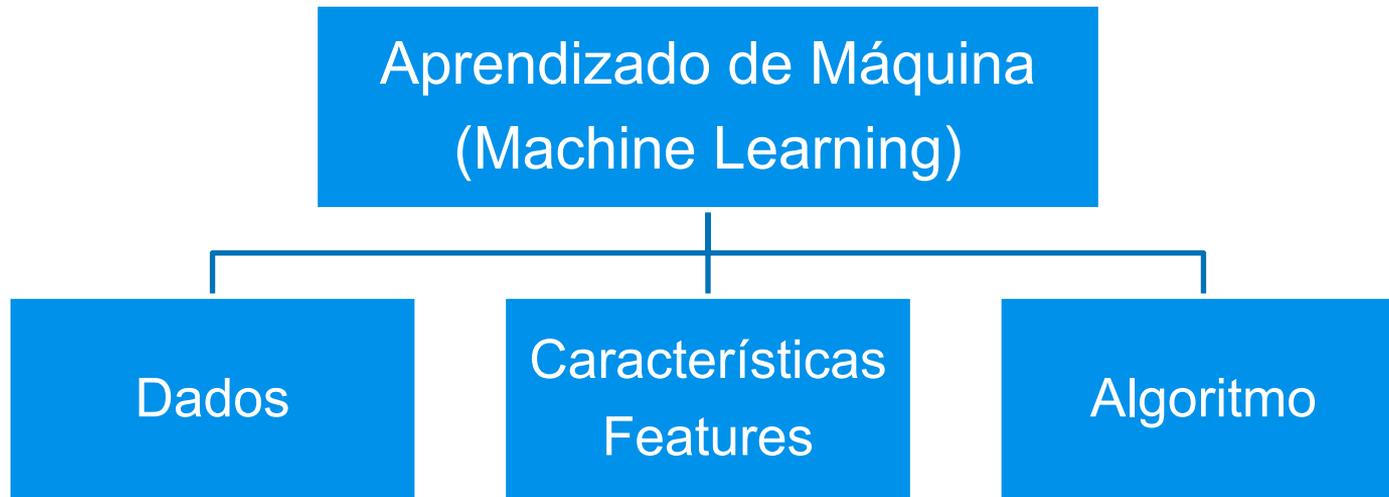
Aprendizado com Estatística

- ◎ Ensinar o computador (Machine Learn – Aprendizado de Máquina) a reconhecer/diferenciar objetos de uma determinada área.
 - Doenças relacionadas a determinados órgãos do corpo humano;
 - Qualidade de diversos produtos (desde roupas a livros);
 - Estratégias de jogo no xadrez.

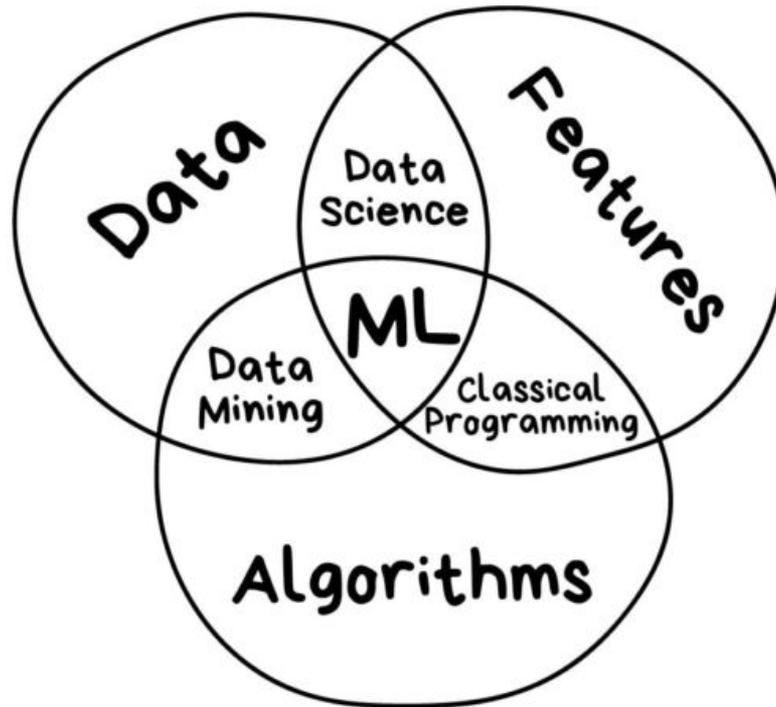
Componentes do aprendizado de máquina

- ◎ O único objetivo do aprendizado de máquina é prever os resultados com base nos dados recebidos;
- ◎ Quanto maior a variedade de amostras, mais fácil será encontrar padrões relevantes para prever um resultado.

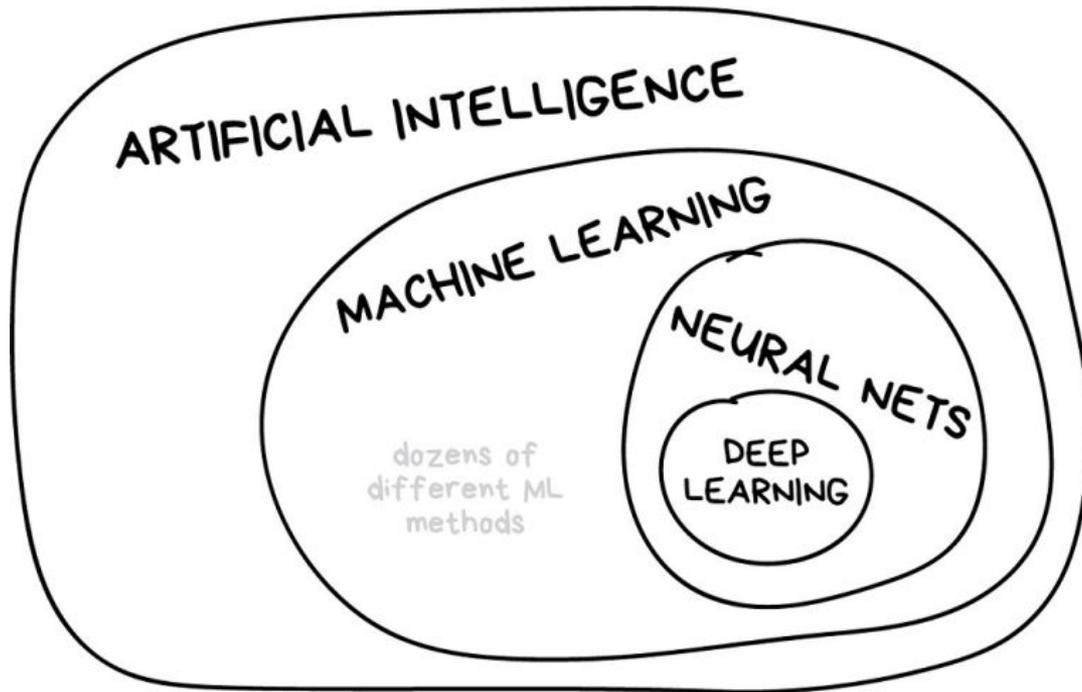
Componentes do aprendizado de máquina



Componentes do aprendizado de máquina



Aprendizagem vs Inteligência



Aprendizagem vs Inteligência

- ◎ **Inteligência artificial** é o nome de todo um campo de conhecimento, semelhante à biologia ou à química.
- ◎ **Aprendizado de Máquina** é uma parte da inteligência artificial. Uma parte importante, mas não a única.
- ◎ **Redes neurais** são um dos tipos de aprendizado de máquina. Um tipo muito popular, mas existem outras ótimas abordagens nessa classe.
- ◎ **Deep Learning** é um método moderno de construção, treinamento e uso de redes neurais.

Aprendizagem vs Inteligência

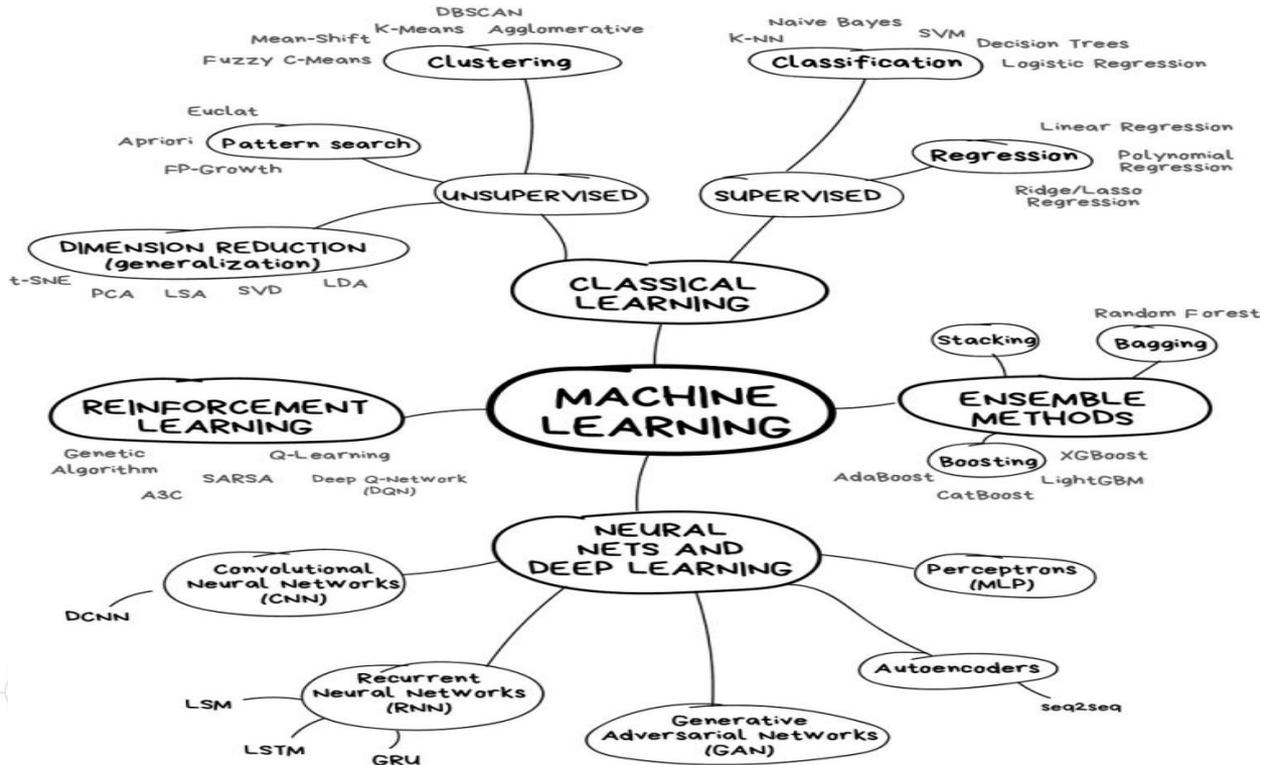
🎯 **Maquinas podem:**

- Realizar previsões;
- Memorizar;
- Reproduzir padrões;
- Escolher o melhor item.

🎯 **Máquinas não podem**

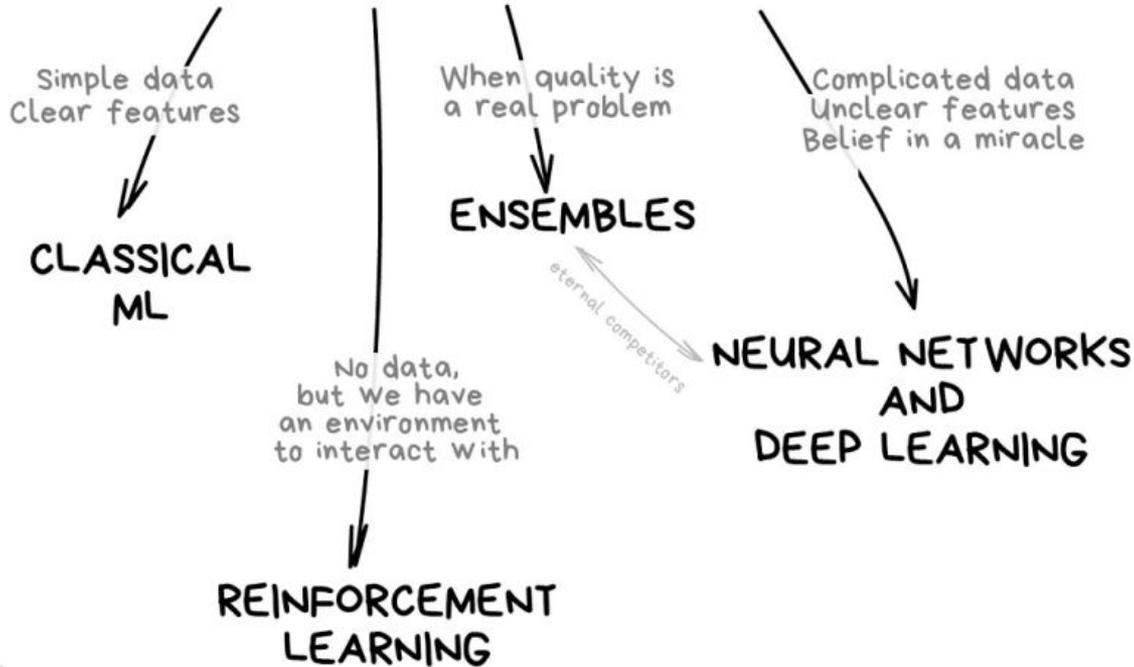
- Criar algo inteiramente novo;
- Ficar muito inteligente rapidamente;
- Ir além das suas tarefas;
- Matar todos os humanos.

O mapa da aprendizagem de máquina

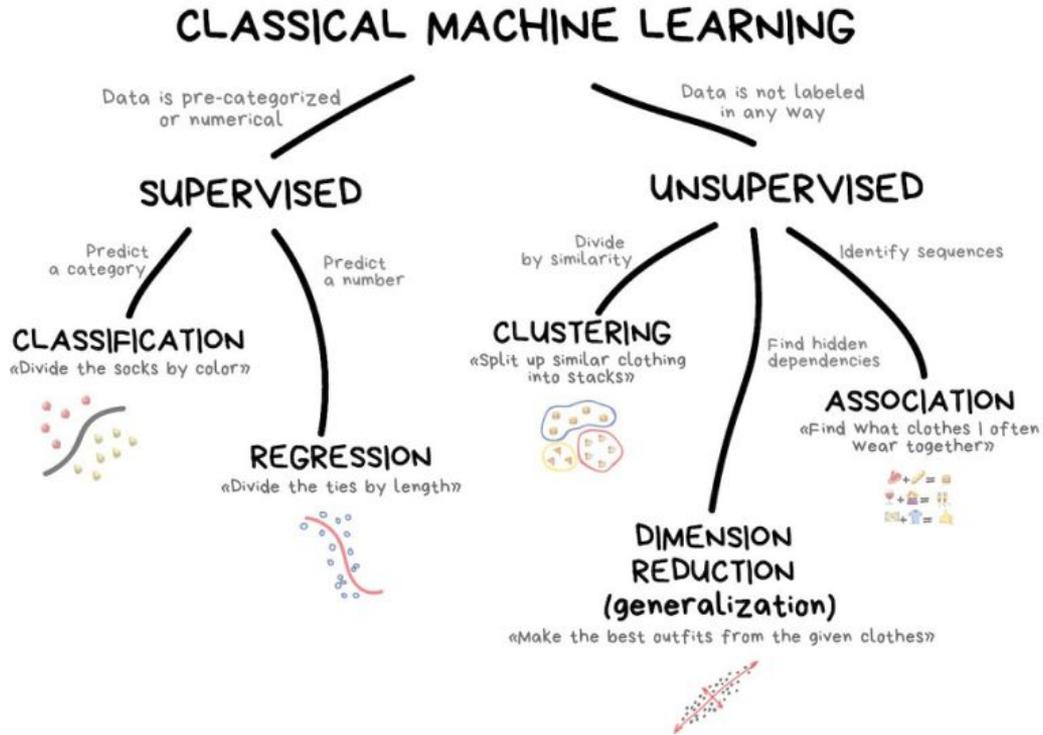


O mapa da aprendizagem de máquina

THE MAIN TYPES OF MACHINE LEARNING



O mapa da aprendizagem de máquina



O mapa da aprendizagem de máquina

⊙ No Aprendizado Supervisionado

- a máquina tem um “supervisor”, que dá à máquina todas as respostas.

⊙ Aprendizado não supervisionado

- significa que a máquina é deixada sozinha com uma pilha de fotos de animais e uma tarefa: descobrir quem é quem.
 - ⊙ Os dados não são rotulados e não há “supervisor”,
 - ⊙ a máquina está tentando encontrar padrões por conta própria.

O mapa da aprendizagem de máquina



- ◎ Aprendizagem Supervisionada
 - Limites Teóricos da Aprendizagem
 - Depende da Teoria da Aprendizagem Estatística.
- ◎ Aprendizagem Não Supervisionada
 - Ainda Carece de Limites Teóricos



Teoria da aprendizagem estatística

- ◎ Sua ideia básica era provar como algum algoritmo supervisionado "aprende" sobre certas condições;
 - Trata-se de encontrar garantias para tarefas de classificação e quantificar quantos dados são necessários para obter tais garantias de aprendizagem.

Teoria da aprendizagem estatística

- ◎ Duas partes principais são fundamentais para o estudo desta teoria de aprendizagem:
 - (1) princípio de Minimização de Risco Empírico (ERM), que aproxima a função de perda verdadeira (e desconhecida) tomando apenas os exemplos observados (do conjunto de treinamento) , e usa a hipótese que minimiza o erro dentro do conjunto de treinamento;
 - (2) o segundo envolve encontrar uma compensação entre a complexidade do espaço de hipóteses, ou seja, o viés do algoritmo de classificação e o erro de classificação calculado nos dados de treinamento (o risco empírico).

Teoria da aprendizagem estatística

- ◎ Premissas necessárias para garantir a aprendizagem no contexto da Teoria da Aprendizagem Estatística, conforme proposto por Vapnik.
 1. Nenhuma suposição é feita sobre a função de probabilidade conjunta $P(X \times Y)$;
 2. Os exemplos devem ser amostrados de $P(X \times Y)$ de maneira independente;
 3. As etiquetas podem assumir valores não determinísticos devido ao ruído e sobreposição de classes;

Teoria da aprendizagem estatística

- ◎ Premissas necessárias para garantir a aprendizagem no contexto da Teoria da Aprendizagem Estatística, conforme proposto por Vapnik.
 4. $P(X \times Y)$ é fixo (estático, portanto não muda ao longo do tempo);
 5. $P(X \times Y)$ é desconhecido no estágio de treinamento.



Estatística e Algoritmos na Educação ?

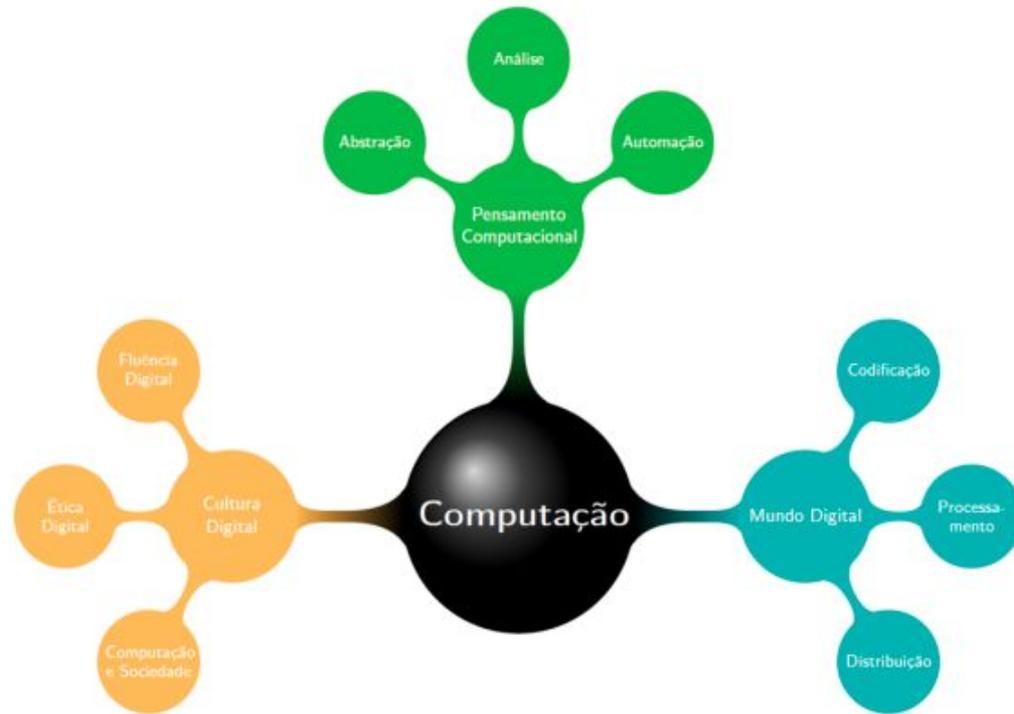
A era dos dados

- ⦿ Esse desenvolvimento rápido das tecnologias digitais, impulsionam a capacidade de criação para além do mundo físico e vêm interferindo na maneira como as pessoas se relacionam, se comunicam e aprendem.
- ⦿ De acordo com a Sociedade Brasileira da computação (SBC), a revolução digital exige um pleno desenvolvimento das habilidades na utilização da tecnologia, ou seja, através do conhecimentos da área de Computação.

Organização do conhecimento

- ◎ Uma das formas de organizar os conhecimentos da área de Computação pode ser estruturada em 3 eixos:
 - Pensamento Computacional,
 - Mundo Digital; e
 - Cultura Digital.

Área de Computação



Pensamento Computacional

- ◎ O Pensamento Computacional se refere à capacidade de sistematizar, representar, analisar e resolver problemas. Apesar de ser um termo recente, vem sendo considerado como um dos pilares fundamentais do intelecto humano, junto com leitura, escrita e aritmética, pois como estes, serve para descrever, explicar e modelar o universo e seus processos complexos.

Pensamento Computacional



- ◎ Selby, Dorling e Woollard (2014) mostram como resultado do seu trabalho que um currículo articulado em que o pensamento computacional seja desenvolvido de forma adequada conduz a uma maior responsabilidade do sujeito em relação ao seu próprio aprendizado, o que faz com que ele aprenda conceitos para toda vida.
- 

Os pilares do Pensamento Computacional

Abstração

Compreende as abstrações necessárias representar informações e processos e as técnicas de construção de soluções (algoritmos).

Análise

Consiste de técnicas de análise de algoritmos quanto a sua correção e eficiência, sob diferentes aspectos.

Automação

Envolve a mecanização das soluções (ou de suas partes), permitindo que máquinas nos ajudem a solucionar os problemas.

Os pilares do Pensamento Computacional



Pensamento Computacional



- ◎ Wing (Wing, 2006) utiliza o termo Pensamento Computacional para apresentar a visão de que todas as pessoas podem se beneficiar do ato de pensar como um cientista da Computação

Pensamento Computacional



- ◎ O desenvolvimento do Pensamento Computacional não tem como objetivo direcionar as pessoas a pensarem como computadores.
- ◎ Ao contrário, sugere que utilizemos a nossa inteligência, os fundamentos e os recursos da computação para abordar os problemas.

Pensamento Computacional

- ◎ Wing (2008) trata a abstração como a essência do pensamento computacional sendo um dispositivo necessário no processo de solução de problemas.
- ◎ A SBC (2018) apresenta como pilar principal da solução de problemas a abstração posto que para que um problema seja resolvido é preciso a construção de um modelo abstrato da realidade.

A estatística e a resolução de problemas

- ◎ A resolução de problemas compreende a utilização de conhecimentos obtidos anteriormente em novas situações;
- ◎ A resolução de problemas pode ser abordada com os conhecimentos relacionados ao pensamento matemático e ao pensamento computacional.

A estatística e a resolução de problemas

- Os algoritmos são modelos que nos ajudam a entender o problema e suas condições de contorno, sendo a projeção e demonstração do processo de solução de problema a ser executado em um computador usando uma estrutura de controle.

A estatística e a resolução de problemas

- ◎ A linguagem matemática se conecta com a apresentação abstrata essencial área de algoritmos;
- ◎ Toda a construção de algoritmos envolve o desenvolvimento de habilidades matemáticas relacionadas com a resolução de problemas;

Planejar e pensar o mundo

- ◎ O pensamento computacional na educação é um conceito de alto impacto e grande poder transformador cada vez mais popular no meio escolar.
- ◎ Essa nova forma de lidar, de analisar, de planejar e de pensar o mundo tende à onipresença na Educação e em todos os aspectos da vida.

Resolução de problemas

- ◎ A habilidade e lógica utilizadas para a resolução de problemas é uma competência fundamental para o futuro;
- ◎ Na Base Nacional Curricular Comum (BNCC) brasileira, o pensamento computacional e a programação aparecem como um esforço na busca de eficazes estratégias e práticas pedagógicas que levem ao desenvolvimento pleno.

Ler ou interpretar...

- ⊙ Quando o indivíduo adquire a capacidade de interpretar as informações de um texto, consegue compreendê-las de forma contextual e as transformar em ferramentas que o auxiliem de forma sistêmica em sua vida, cumpre-se o objetivo do desenvolvimento pleno.

Inclusão digital



“Compreender, utilizar e criar tecnologias digitais de informação e comunicação de forma crítica, significativa, reflexiva e ética nas diversas práticas sociais (incluindo as escolares) para se comunicar, acessar e disseminar informações, produzir conhecimentos, resolver problemas e exercer protagonismo e autoria na vida pessoal e coletiva.” (BNCC, 2018)



“A pesquisa científica é um processo de aprendizado dirigido. O objetivo dos métodos estatísticos é tornar este processo o mais eficiente possível”.



Obrigado!

Perguntas ?



fb.com/gesielrios



[@gesielrios](https://www.instagram.com/gesielrios)



[@gesielrios](https://twitter.com/gesielrios)



github.com/gesielrios



gesielrios@usp.br

Referências

- ◎ AHO, Alfred V. Computation and computational thinking. The computer journal, v. 55, n. 7, p. 832-835, 2012.
- ◎ ARAUJO, Ana Liz; ANDRADE, Wilkerson; GUERRERO, Dalton. Um mapeamento sistemático sobre a avaliação do pensamento computacional no Brasil. In: Anais dos Workshops do Congresso Brasileiro de Informática na Educação. 2016. p. 1147.
- ◎ BARBOSA, Luciana. A inserção do Pensamento Computacional na Base Nacional Comum Curricular: reflexões acerca das implicações para a formação inicial dos professores de matemática. In: Anais do XXV Workshop de Informática na Escola. SBC, 2019. p. 889-898.

Referências

- ◎ BARCELOS, Thiago Schumacher; SILVEIRA, Ismar Frango. Pensamento computacional e educação matemática: Relações para o ensino de computação na educação básica. In: XX Workshop sobre Educação em Computação, Curitiba. Anais do XXXII CSBC. 2012. p. 23.
- ◎ BOSCHETTI, Alberto; MASSARON, Luca. Python data science essentials. Packt Publishing Ltd, 2015.
- ◎ DA SILVA BARBOSA, Luciana Leal; MALTEMPI, Marcus Vinícius. Matemática, Pensamento Computacional e BNCC: desafios e potencialidades dos projetos de ensino e das tecnologias na formação inicial de professores. Revista Brasileira de Ensino de Ciências e Matemática, v. 3, n. 3, 2020.

Referências

- ◎ HASLWANTER, Thomas. An Introduction to Statistics with Python. With applications in the life sciences. Switzerland: Springer International Publishing, 2016.
- ◎ JAMES, Gareth et al. An introduction to statistical learning. New York: springer, 2013.
- ◎ JÚNIOR, J. C. R. P. et al. Ensino de algoritmos e programação: uma experiência no nível médio. In: XIII Workshop de Educação em Computação (WEI'2005). São Leopoldo, RS, Brasil. 2005.
- ◎ LOPES, Gesiel Rios et al. Introdução à Análise Exploratória de Dados com Python.

Referências

- ◎ MELLO, Rodrigo F.; PONTI, Moacir Antonelli. Machine learning: a practical approach on the statistical learning theory. Springer, 2018.
- ◎ MESTRE, Palloma Alencar Alves et al. O uso do pensamento computacional como estratégia para resolução de problemas matemáticos. 2017.
- ◎ SAVIANI, Dermeval. Educação escolar, currículo e sociedade: o problema da Base Nacional Comum Curricular. Movimento-revista de educação, n. 4, 2016.
- ◎ SILVA, Lenildes Ribeiro. Unesco: Os quatro pilares da “educação pós-moderna. Revista Inter Ação, v. 33, n. 2, p. 359-378, 2008.

Referências

- ◎ TUKEY, John W. et al. Exploratory data analysis. 1977.
- ◎ VAPNIK, Vladimir N. An overview of statistical learning theory. IEEE transactions on neural networks, v. 10, n. 5, p. 988-999, 1999.
- ◎ VAPNIK, Vladimir. The nature of statistical learning theory. Springer science & business media, 2013.
- ◎ WING, Jeannette M. Computational thinking and thinking about computing. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, v. 366, n. 1881, p. 3717-3725, 2008.
- ◎ WING, Jeannette M. Computational thinking. Communications of the ACM, v. 49, n. 3, p. 33-35, 2006.



That's all Folks!