

Data Science e Machine Learning na Prática: Introdução e Aplicações na Indústria de Processos

Apresentação 2 - Perspectivas históricas

Afrânio Melo

afraeq@gmail.com

afrrjr.weebly.com

Escola Piloto Prof. Giulio Massarani
PEQ-COPPE-UFRJ

2019

A Quarta Revolução Industrial - Indústria 4.0

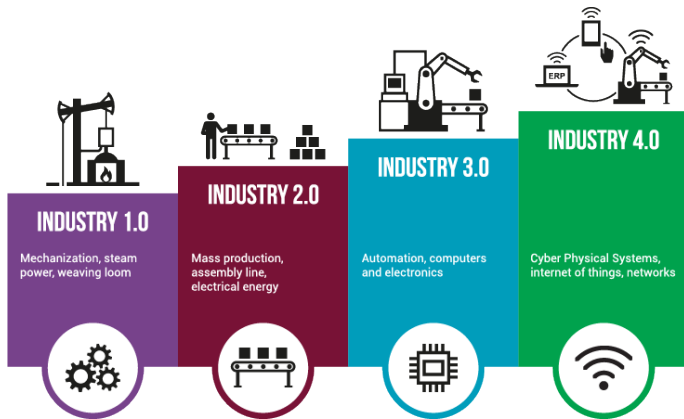


Figura 1: Evolução das revoluções industriais.

A Quarta Revolução Industrial - Indústria 4.0

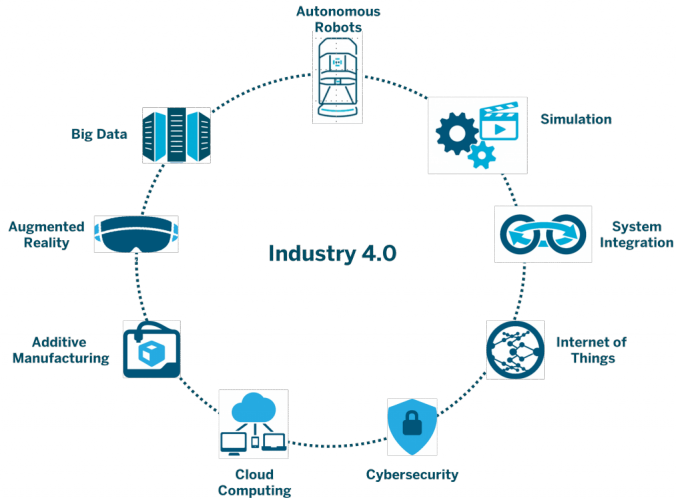


Figura 2: Pilares da Indústria 4.0. Será que estamos prontos para a mudança?

Data Science, Machine Learning e a Engenharia de Processos



Figura 3: Venkat Venkatasubramanian.

- Nesta seção, nosso objetivo é:
 - entender como o crescimento dos campos de Data Science e Machine Learning e, mais genericamente, a Inteligência Artificial impactaram, impactam e provavelmente impactarão o campo da Engenharia de Processos.
- Esta seção é baseada em um trabalho recente publicado pelo engenheiro indiano Venkat Venkatasubramanian ^[1].

[1] Venkat Venkatasubramanian. "The promise of artificial intelligence in chemical engineering: Is it here, finally?"
Em: *AIChE Journal* 65.2 (2019), pp. 466–478. DOI: 10.1002/aic.16489.

Data Science, Machine Learning e a Engenharia de Processos

E a engenharia de processos?

- É evidente que os termos *data science* e *machine learning* estão na moda!
- Também é evidente o fato de essas áreas do conhecimento estarem influenciando de forma decisiva o evoluir de nossa sociedade, principalmente por meio de impactos nos campos da computação, da eletrônica e dos sistemas de informação.
- Mas qual o impacto do *data science* e do *machine learning* no campo da engenharia de processos? Qual a história desse impacto até o momento e quais as perspectivas para o futuro?

Data Science, Machine Learning e a Engenharia de Processos

As 3 fases da AI na engenharia de processos

- O estudo da inteligência artificial (*artificial intelligence*, AI) na engenharia de processos pode ser descrito em três fases:
 - **Fase I:** sistemas especialistas;
 - **Fase II:** redes neurais;
 - **Fase III:** *deep learning* e *data science*.

Data Science, Machine Learning e a Engenharia de Processos

Fase I: sistemas especialistas (~1983 à ~1995)

- *Sistemas especialistas* são programas de computador que mimetizam o método de resolução de problemas dos seres humanos utilizando conhecimento do domínio acumulado ao longo da experiência humana (possivelmente na forma de *heurísticas*).
- Nesses sistemas, a ordem de execução é separada do conhecimento do domínio! Isso proporciona flexibilidade e a possibilidade de adição de conhecimento de forma incremental.

Data Science, Machine Learning e a Engenharia de Processos

Fase I: sistemas especialistas (~1983 à ~1995)

- Vantagens:
 - representação e arquitetura intuitivas;
 - há explicações plausíveis para as decisões tomadas pelo sistema.
- Desvantagens:
 - alto custo (dinheiro, tempo e esforço) para implementar, manter e monitorar um sistema especialista confiável;
 - dificuldade de escalar o sistema.
- Algumas aplicações:
 - predição de propriedades termofísicas de misturas complexas de fluidos;
 - *design* de catalisadores;
 - síntese de processos.

Data Science, Machine Learning e a Engenharia de Processos

Fase II: redes neurais (~1990 à ~2008)

- Primeira aplicação séria do *machine learning*: o conhecimento agora vem da máquina aprendendo a partir dos dados e não da experiência humana.
- Tudo começou com a proposta de um algoritmo de retropropagação para o treinamento de redes neurais *feedforward* em 1986 (Rumelhart, Hinton e Williams [2]) com o objetivo de se aprender padrões ocultos em dados de entrada-saída.

[2] David E. Rumelhart, Geoffrey E. Hinton e Ronald J. Williams. "Learning representations by back-propagating errors". Em: *Nature* 323.6088 (1986), pp. 533–536. DOI: 10.1038/323533a0. URL: <https://doi.org/10.1038/323533a0>.

Data Science, Machine Learning e a Engenharia de Processos

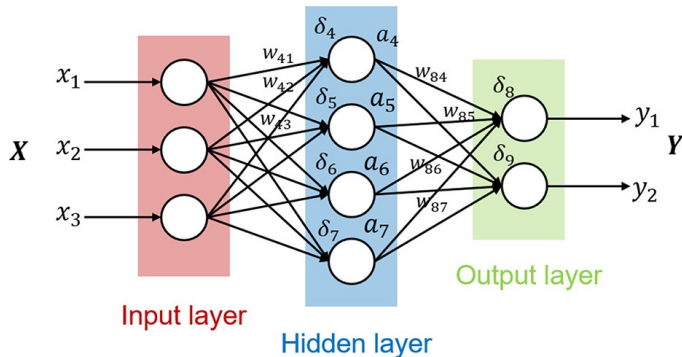


Figura 4: Arquitetura de uma rede *feedforward*.

Data Science, Machine Learning e a Engenharia de Processos

Fase II: redes neurais (~1990 à ~2008)

- Esse novo algoritmo fez com que problemas fortemente não-lineares pudessem ser modelados de forma automática e simples!
- Surgiu então uma imensa quantidade de trabalhos usando redes neurais em várias áreas do conhecimento! Especificamente na área de engenharia de processos, muito se avançou em termos principalmente de modelagem e controle de processos e detecção e diagnóstico de falhas.

Data Science, Machine Learning e a Engenharia de Processos

Fase II: redes neurais (~1990 à ~2008)

- Mas, logo depois do começo, as coisas esfriaram e pararam de avançar.
- Problemas mais complexos surgiam e as redes não eram capazes de resolver.
- Suspeita: necessidade de cada vez mais camadas ocultas, além de apenas uma!
- Mas, na época, parecia simplesmente impossível treinar esse tipo de rede...

Data Science, Machine Learning e a Engenharia de Processos

Falta de impacto nas fases I e II

- Apesar dos avanços, não houve impacto real da inteligência artificial na indústria de processos, nas fases I e II! Vários são os motivos:
 - limitação de dados;
 - limitação na capacidade computacional;
 - falta de interesse do mercado, já que abordagens mais tradicionais (MPC, otimização, etc.) evoluíam mais rápido e forneciam resultados mais satisfatórios.
- Não havia o *empurrão da tecnologia* e nem o *puxão do mercado*!

Data Science, Machine Learning e a Engenharia de Processos

Heureka!

- Até que, em 2006, veio o *breakthrough*: Hinton GE, Osindero S e Teh Y-W.^[3] publicaram uma metodologia para treinar redes neurais com várias camadas em um intervalo de tempo factível!
- Esse tipo de abordagem hoje é conhecido como *deep learning*.
- A proposta revolucionou a inteligência artificial, aumentando incrivelmente a capacidade de treinamento de algoritmos de aprendizado e tornando possível todos os avanços que testemunhamos hoje.

[3] Geoffrey E. Hinton, Simon Osindero e Yee-Whye Teh. "A Fast Learning Algorithm for Deep Belief Nets". *Em: Neural Comput.* 18.7 (jul. de 2006), pp. 1527–1554. ISSN: 0899-7667. DOI: 10.1162/neco.2006.18.7.1527. URL: <http://dx.doi.org/10.1162/neco.2006.18.7.1527>.

Data Science, Machine Learning e a Engenharia de Processos

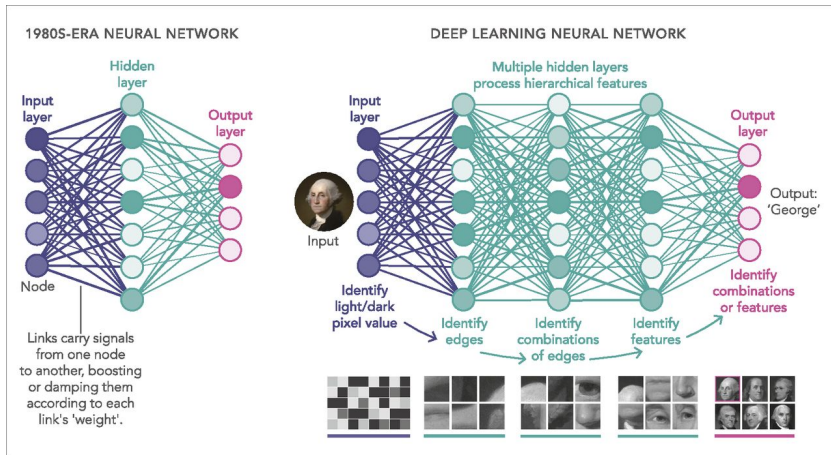


Figura 5: Comparação entre as arquitetura “tradicional” e “profunda” das redes neurais.

Data Science, Machine Learning e a Engenharia de Processos

Fase III: *data science* e *deep learning* (~2005 à ~hoje)

- Foi o ponto de partida para entrarmos na fase atual.
- É extensivo hoje o uso de métodos de *deep learning*, aprendizado por reforço e *machine learning* com forte base estatística, tendo ênfase metodologia *data science*.

Data Science, Machine Learning e a Engenharia de Processos

As coisas são diferentes na fase III?

- Alguns motivos pelos quais podemos supor que o cenário na fase III é diferente:
 - a tendência da indústria 4.0, com mais instrumentação e sensores, devem tornar os dados cada vez mais abundantes;
 - a capacidade computacional só faz aumentar (a lei de Moore ainda não falhou!);
 - o interesse do mercado é cada vez maior, por conta do sucesso da AI em outras áreas e pelo fato das abordagens tradicionais (como MPC) já terem fornecido a maior parte dos benefícios de que são capazes, diminuindo a capacidade de evolução adicional;
 - tipicamente, uma tecnologia demora 50 anos para amadurecer, penetrar e provocar um impacto generalizado, desde a descoberta até a adoção final (Aspen Plus e MPC, por exemplo). Isso indica que a AI deve atingir a plenitude do seu impacto na década de 2030.
- Agora parece haver o *empurrão da tecnologia e puxão do mercado!!*

Data Science, Machine Learning e a Engenharia de Processos

Um banho de água fria

- Apesar de haver motivo para animação, também há para desconfiança:
 - ao contrário dos campos de finanças, visão, linguagem, redes sociais, etc., nosso campo não é verdadeiramente *big data*, de acordo com a definição de *big data* apresentada na primeira aula;
 - o *deep learning* fornece modelos cegos, que não permitem analisar os motivos de suas tomadas de decisão, tornando-os menos propensos à intuição e interpretação humanas;
 - o *deep learning* é, em última análise, um método ineficiente. Uma criança precisa de um ou dois exemplos para aprender a reconhecer, digamos, uma vaca, enquanto um algoritmo de *deep learning* precisa treinar milhares de vezes com milhares de vacas;
 - as redes neurais humanas nascem com instintos, intuição, acumulam memória de toda uma vida, etc. As redes artificiais atuais, em geral, começam do zero. Como mimetizar todo o acúmulo de informação que ocorre no cérebro humano?
 - E o eterno problema da inteligência artificial: a falta de *consciência* (seria isso bom ou ruim?).

Data Science, Machine Learning e a Engenharia de Processos

Desafios para a evolução da AI na engenharia de processos

- avançar na melhora da instrumentação industrial, de modo a aumentar a quantidade de dados disponíveis;
- avançar além da abordagem puramente centrada em dados, explorando outros aspectos da AI, como o raciocínio baseado em estruturas e relações simbólicas;
- desenvolvimento de modelos híbridos, baseados tanto em dados quanto no conhecimento tradicional da engenharia de processos (termodinâmica, cinética, fenômenos de transporte, etc.);
- aplicar a AI para o desenvolvimento de *teorias de emergência*, talvez o problema científico mais desafiador do séc XXI, que em tese se solucionado poderia resolver a *questão da natureza da consciência*;
- adaptar o ensino nas universidades para que reflita essa nova classe de conhecimentos.

Data Science, Machine Learning e a Democracia Liberal

- Nesta seção, nosso objetivo é:
 - utilizar alguns conceitos de História e relacioná-los com o crescimento dos campos de Data Science e Machine Learning;
 - demonstrar assim como esses campos serão essenciais para a evolução da sociedade no futuro próximo.
- Esta seção é baseada em um trabalho recente publicado pelo historiador israelense Yuval Noah Harari [4].



Figura 6: Yuval Noah Harari.

[4] Yuval Noah Harari. *Homo Deus - A Brief History of Tomorrow*. Harper, 2017.

Um pouquinho de História...



Por que os seres humanos controlam o planeta?

- É evidente que os seres humanos controlam o planeta inteiro.
- A pergunta é: por quê? Como isso aconteceu?

Um pouquinho de História...

Por que os seres humanos controlam o planeta?

- Se você respondeu "*por que somos mais inteligentes*", lembre-se de nos nossos ancestrais pré-históricos, há 70.000 anos atrás.
- A maioria deles era muito mais inteligente do que nós (individualmente), e mesmo assim a humanidade não controlava o planeta!
- O impacto dos seres humanos no mundo na verdade era insignificante.



Um pouquinho de História...



Por que os seres humanos controlam o planeta?

- O segredo do nosso sucesso está na capacidade de *cooperação*.
- Os seres humanos são os únicos animais capazes de *cooperar em grande número e de forma flexível*.
- Formigas conseguem cooperar em grande número, mas não de forma flexível. Chimpanzés conseguem cooperar de forma flexível, mas não em grande número. Apenas os humanos conseguem os dois!!

Um pouquinho de História...

Por que os seres humanos controlam o planeta?

- Pergunta: por que os seres humanos conseguem cooperar de forma tão eficiente? Como exatamente isso acontece?

Um pouquinho de História...

Por que os seres humanos controlam o planeta?

- Resposta: por conta da nossa imaginação!
- Os seres humanos são os únicos animais capazes de inventar histórias (ficções). Se todos acreditam nelas, a cooperação se torna possível.
- Exemplos de algumas histórias que permitem cooperação em larga escala: Brasil, Estados Unidos, União Soviética, Toyota, Google, direitos humanos, mercado financeiro, Código Penal, Ministério Público, Banco Mundial, ONU, real, dólar (essa última sendo a mais poderosa de todas).
- Os animais vivem em duas realidades: objetiva e subjetiva. Seres humanos criam um novo tipo de realidade: a *realidade inter-subjetiva*.
- Os demais animais usam a comunicação para descrever as realidades objetiva e subjetiva, e não para estendê-la, como nós.

Um pouquinho de História...

Humanismo: a história dominante

- A história dominante dos últimos séculos foi a história do *humanismo*: a crença de que o *Homo sapiens* tem uma natureza sagrada, diferente de todos os demais fenômenos e seres do mundo natural.
- Essa história tem três vertentes principais:
 - 1) humanismo liberal:
 - 2) humanismo socialista:
 - 3) humanismo evolucionário.

	Humanismo liberal	Humanismo socialista	Humanismo evolucionário
	<i>A espécie Homo sapiens tem uma natureza sagrada, especial em relação a todos os demais fenômenos e seres.</i>		
O que é a humanidade?	É uma característica de cada Homo sapiens individual.	É uma característica coletiva e reside na espécie como um todo.	É uma característica mutável que pode degenerar ou evoluir.
Qual o mandamento supremo?	Proteger as liberdades individuais, os direitos humanos, a democracia, o livre arbítrio e o livre mercado.	Proteger a igualdade entre os membros da espécie.	Proteger a humanidade de degenerar em sub-humanos e estimular a evolução para super-humanos.

Um pouquinho de História...

Humanismo liberal: a história vitoriosa

- **Segunda Guerra Mundial:** confronto entre essas três vertentes do humanismo, com derrota do humanismo evolucionário!
- **Guerra Fria:** confronto entre as duas vertentes restantes, com derrota do humanismo socialista!

Um pouquinho de História...

Humanismo liberal: a história vitoriosa (até quando?)

- Portanto, a história vitoriosa e dominante de trinta anos para cá é o humanismo liberal.
- A grande questão é que a fé nessa história vem diminuindo. Há uma percepção cada vez maior de que ela não é capaz de resolver nossos problemas.
- Provavelmente vivemos um período de transição muito especial, em que a história dominante está prestes a mudar.
- Pergunta mais importante dos dias atuais: *qual história vai substituir o humanismo liberal?*

Ameaças ao humanismo liberal

Inteligência, sim. Consciência, para quê?

- Um dos principais princípios que ameaçam o humanismo liberal nos próximos anos é:

Para as instituições e empresas, a inteligência é essencial, mas a consciência é opcional.

Ameaças ao humanismo liberal

Inteligência, sim. Consciência, para quê?

- A automação do século passado se baseou na *automação de tarefas mecânicas*.
- A automação do século XXI se baseia na *automação do reconhecimento de padrões*.
- Tente pensar em algumas profissões que se baseiam em reconhecimento de padrões.

Ameaças ao humanismo liberal

Inteligência, sim. Consciência, para quê?

- Exemplos: motoristas, contadores, investidores da bolsa, enxadristas, médicos, farmacêuticos, advogados, músicos...
- Perigo: o surgimento de uma classe de pessoas economicamente inúteis!
- Uma das principais perguntas do séc XXI provavelmente será: *o que fazer com essas pessoas?*

Ameaças ao humanismo liberal

A autoridade passa para os algoritmos

- Outro princípio que ameaça o humanismo liberal:

*Seres humanos tomam decisões de acordo com algoritmos bioquímicos.
Algoritmos artificiais são mais sofisticados, eficientes e podem trabalhar em rede.
Portanto, são capazes de tomar melhores decisões.*

Ameaças ao humanismo liberal

A autoridade passa para os algoritmos

- Se os algoritmos eletrônicos são capazes de tomar as melhores decisões, por que não deixá-los fazer isso?
- Achou assustador? Pensa que nunca vai acontecer? Então reflita sobre esses casos:
 - sistemas de navegação GPS;
 - sistemas de recomendação;
 - sistemas de busca online (quem passa da primeira página do Google?).
- Há uma ilusão bastante difundida de que as pessoas prezam muito o ato de tomar decisões. Na verdade, as pessoas prezam suas *certezas*. Decidir perante à incerteza é angustiante e pode causar grande sofrimento.

Ameaças ao humanismo liberal

A autoridade passa para os algoritmos

- Há dois mecanismos principais pelos quais os algoritmos podem aumentar sua capacidade de tomar decisões sobre nós:
 - 1) pela obtenção, em rede, de uma grande quantidade de dados da sociedade como um todo;
 - 2) pelo monitoramento cada vez mais completo do indivíduo por meio de sensores bioquímicos e *gadgets* biônicos.
- Exemplos que já acontecem: detecção de uma epidemia por meio de estatísticas de buscas no Google; monitoramento da velocidade de leitura de diferentes páginas de um livro no Kindle.
- Exemplos que provavelmente acontecerão: detecção de doenças individuais por meio do monitoramento da bioquímica do corpo; detecção de sentimentos e emoções por meio do monitoramento dessa mesma bioquímica.

Ameaças ao humanismo liberal

A autoridade passa para os algoritmos

- Uma vez que os algoritmos passam a saber cada vez mais sobre nós (graças às redes sociais, aos sensores bioquímicos e aos dispositivos biônicos), e sobre a sociedade como um todo (já que trabalham em rede), torna-se cada vez mais confortável deixar as decisões por sua conta.
- Tanto as decisões que seriam tomadas por outros seres humanos (ex: tratamentos médicos) quanto as decisões pessoais (ex: que música ouvir? Quem escolher para governar? Qual faculdade cursar? Com quem se casar?) seriam tomadas de forma muito mais simples e, o mais importante, *com maior probabilidade de acerto*.
- Seria o fim do livre arbítrio e da suprema autoridade do indivíduo para fazer suas escolhas.

Ameaças ao humanismo liberal

A autoridade passa para os algoritmos

- Muitos podem argumentar que, nesse cenário, haveria uma resistência grande a fazermos parte do sistema, em nome de valores como, por exemplo, privacidade.
- Só que na verdade, já fornecemos nossos dados e nossa privacidade a essas empresas em troca de simples serviços de e-mail e vídeos de gatinhos.
- Imagine quando o oferecido na troca for bem mais valioso, como nossa saúde ou bem-estar social?
- Se as coisas continuarem a evoluir nesse passo, chegará um momento em que será impossível se desconectar da rede. Desconexão significará morte. Ex: um mundo em que sensores biométricos e nano-robôs online se integrem ao corpo humano.

Dataísmo - uma nova história dominante

Dataísmo

- O historiador Yuval Noah Harari tem um nome para essa nova possível história dominante, em que a autoridade passa do indivíduo para os algoritmos: *dataísmo* (do inglês *data*, que significa dados).
- No dataísmo, o universo é percebido como um fluxo constante de dados e os seres humanos são elementos nesse fluxo.

Dataísmo - uma nova história dominante

Dataísmo

- Capitalistas liberais acreditam na “mão invisível do mercado”. Dataístas acreditam na “mão invisível do fluxo de dados”.
- O dataísta encara a Quinta Sinfonia de Beethoven, uma bolha no mercado financeiro e o vírus da gripe suína como três padrões de fluxo de dados que podem ser analisados utilizando os mesmos conceitos e ferramentas.

Dataísmo - uma nova história dominante

Dataísmo

- Organismos, ecossistemas, sociedades, economias e estruturas políticas podem ser interpretadas como sistemas de aquisição, processamento e análise de dados!
- Quanto mais eficiente o sistema, maior sua vantagem competitiva em relação aos demais.
- Essa visão pode explicar, por exemplo:
 - a evolução biológica pela seleção natural;
 - a vitória do capitalismo sobre o comunismo;
 - a vitória das ditaduras sobre as democracias na maior parte da História;
 - a vitória das democracias sobre as ditaduras nas últimas décadas;
 - declínio da fé nas democracias nos últimos anos!

Dataísmo - uma nova história dominante

Dataísmo

- Diria o dataísta: “Se você experimentou algo, grave. Se gravou, faça upload. Se fez upload, compartilhe”.
- O supremo valor está na conexão ao fluxo de dados!

Dataísmo - uma nova história dominante

Dataísmo

- Muito ainda se discute sobre se ditaduras militares, comunismo, fascismo, etc, mas será que essa discussão não está se tornando obsoleta? Será que o mundo está se encaminhando para ser dataísta, sem nem se dar conta?
- Se você acha tudo isso muito louco, pense em como você se sente quando sua conexão à Internet é interrompida à sua revelia. Qual o valor que *você* atribui à conexão ao fluxo de dados?
- Seria você um dataísta??

Obrigado pela atenção!