

Disciplina de Módulo II: Teórico Avançado

Disciplina: P07561 - Fundamentos da Inteligência Artificial (FIA)

Nível: Mestrado e Doutorado

Area de Concentração: Processos Cognitivos e Ambientes Digitais

Linha de Pesquisa: Inteligência Artificial e Gestão

Professor: Dra. Dora Kaufman (cód. 307054; responsável)

Horário: 2ª feira, das 9h00 às 12h00

Créditos: 3

Carga Horária: 255 horas

1. Descrição e ementa da disciplina

O objetivo da disciplina é explorar os fundamentos da inteligência artificial (IA), adentrando nos meandros da própria tecnologia e alguns de seus modelos, e estudos de caso em distintos domínios de implementação. Estudaremos o histórico da formação do campo de conhecimento, a técnica de redes neurais profundas (*deep learning*), as duas categorias - IA preditiva e IA generativa - e as respectivas arquiteturas; os benefícios dos modelos, que como tecnologia de propósito geral estão alterando a lógica de funcionamento da economia e da sociedade; os riscos associados e os potenciais danos; e as iniciativas regulatórias em curso mundo afora e alguns dos framework de governança propostos por instituições e/ou especialistas.

2. Detalhamento da ementa

1a. semana

Apresentação do curso/alunos, metodologia e processo de avaliação. Introdução à disciplina.

2a. semana

Visão geral da inteligência artificial, tecnologia de propósito geral, origem e conteúdo do conceito. Histórico das GPT, ilustrando seus impactos transformadores na economia e na sociedade.

3a. semana

Histórico e Conceituação da Inteligência Artificial - Ficção versus realidade.

4a. semana

Machine learning, redes neurais e deep learning.

5a. semana

Inteligência de máquina vs inteligência humana. Barreiras para atingir a cognição humana: raciocínio contrafactual, intencionalidade, analogia.

6a. semana

IA preditiva ilustrando com casos de uso.

7a. semana

IA preditiva ilustrando com casos de uso.

8a. semana

IA generativa, ilustrando com casos de uso.

9a. semana

IA generativa, ilustrando com casos de uso.

10a. semana

Impactos na economia: novas lógicas, novos modelos de negocio.

11a. semana

Reação da sociedade aos potenciais danos: breve histórico e principais iniciativas.

12a. Semana

Questões éticas no recente avanço da IA. Desafios regulatórios e legais.

13a. semana

Regulamentação Europa, EUA, Brasil.

14a. semana

Framework de governança.

15a. semana

Framework de governança.

16a. semana

Efeitos do avanço da IA no mercado de trabalho: deslocamento pela automação, efeito negativo sobre a renda e qualificação /requalificação.

17a. semana

Revisão geral e encerramento.

3. Bibliografia

3.1. Básica

AUTOR, David; Mindell, David A.; Reynolds, Elisabeth B. Cambridge, Massachusetts: MIT Press, 2023.

BLACKMAN, Reid. Ethical Machines. Boston, Massachutees: Harvard Business Review Press, 2022.

CHRISTIAN, Brian. O Humano mais Humano: O que a inteligência artificial nos ensina sobre a vida. SP: Companhia das Letras, 2013.

_____. The Alignment Problem: Machine Learning and Human Values. NT: W.W. Norton&Company, 2020.

COECKELBERGH, Mark. AI Ethics. Cambridge, MA: The MIT Press, 2020.

DIGNUM, Virginia. Responsible Artificial Intelligence. Sweden: Springer, 2019.

COLEMAN, Foynn. A Human Algorithm: How Artificial Intelligence is Redefining Who We Are. Berkeley, California: Counterpoint, 2019.

COSSIN, Didier. High Performance Boards: Improving and Energizing Your Governance. UK: Wiley, 2020.

FRISCHMANN, Brett; Selinger, Evan. Re-Engineering Humanity. USA: Cambridge University Press, 2018.

GUNKEL, David J. The Machine Question: Critical Perspectives on AI, Robots, and Ethics. Cambridge, Massachussets: The MIT Press, 2012.

KAUFMAN, Dora. A inteligência artificial irá suplantará a inteligência humana?. 1. ed. São Paulo: Estação das Letras e Cores, 2019.

KAUFMAN, Dora. Desmistificando a Inteligência Artificial. São Paulo: Autêntica, 2022.

MITCHELL, Melanie. Artificial Intelligence: A Guide of Thinking Humans. NY: Farrar, Straus and Giroux, 2019.

PEARL, Judea; Mackenzie, Dana. The Book of Why: The New Science of Cause and Effect. NY: Basic Books, 2018.

RUSSERLL, Stuart. Human Compatible: artificial intelligence and the problem of control. US: Viking, 2019.

SANTAELLA, Lucia. Neo-Humano: A sétima revolução cognitiva do Sapiens. SP: Paulus, 2022.

_____. A inteligência artificial é inteligente? SP: Almedina, 2023.

SCHNEIDER, Susan. Artificial You: AI and the Future of Your Mind. Oxford: Princeton University Press, 2019.

SULEYMAN, Mustafa. The Coming Wave: AI, Power and the 21st Century's Greatest Dilemma. UK: Penguin Randon House, 2023.

3.2. Adicional

AGRAWAL, A.; GANS, J.; GOLDFARB, A. Prediction Machines: The Simple Economics of Artificial Intelligence. Massachusetts: Harvard Press, 2018.

ALPAYDIN, E. Machine Learning. Cambridge, MA: MIT Press, 2016.

ARIEL, E.; STUCKE, M. Virtual Competition: The Promise and Perils of the Algorithm-Driven Economy. Cambridge, MA: Harvard University Press, 2016.

CHRISTINAN, Brian; Griffiths, Tom. Algoritmos to Live By: The Computer Science of Human Decisions.

DOMINGOS, Pedro. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake our World. NY: Basic Books, 2015.

FORD, Martin. Architects of Intelligence: the truth about AI from the people building it. UK: Packt, 2018.

GERRISH, Sean. How Smart Machines Think. Cambridge, Massachusetts: The MIT Press, 2018.

GOODFELLOW, I., Y. Bengio, A. Courville. Deep Learning. Cambridge: MIT Press, 2016.

IANSITI, Marco; LAKHANI, Karin. Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World. Cambridge, MA: Harvard University Press, 2020.

KING, Brett. Bank 4.0: Banking Everywhere, Never at a Bank. US: Wiley, 2019.

LEONHARD, Gerd. Technology vs. Humanity: The coming Clash between man and machine. UK: www.techvshman.com, 2016.

MAYER-SCHONBERGER, Viktor; RAMGE, Thomas. Reinventing Capitalism in the Age of Big Data. NY: Basic Books, 2018.

MAYER-SCHONBERGER, Viktor; CUKIER, Kenneth. Big Data: A revolution that will transform how we live, work, and think. NY: Houghton Mifflin Harcourt, 2013.

RUSSELL, S., NORVIG, P. Inteligência Artificial. 3ª ed. Rio de Janeiro: Elsevier, 2013.

PASQUALE, F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2015.

SEJNOWSKI, Terrence. The Deep Learning Revolution. Cambridge, MA: The MIT Press, 2018.

4. Avaliação

Participação e colaboração para os debates em sala de aula, leituras da bibliografia e dos textos sugeridos ao longo da disciplina. Trabalho final: artigo acadêmico para publicação (não precisa ser individual).