

MÓDULO 2 – TEÓRICO AVANÇADO

Disciplina: PO7808 - Desenvolvimento de Soluções de Inteligência Artificial para a Gestão (DSDIAG)

Nível: Mestrado e Doutorado

Modulo: 2

Area de Concentração: Processos Cognitivos e Ambientes Digitais

Linha de Pesquisa: Inteligência Artificial e Gestão

Professor: Dora Kaufman

Semestre: 1º semestre de 2024

Horário: 2ª feira, das 9h00 às 12h00

Créditos: 3

Carga Horária: 255 horas

Tipo: eletiva

1. Descrição do tema e da ementa da disciplina

O objetivo da disciplina é explorar a natureza da inteligência artificial como tecnologia de propósito geral (GPT, *general purpose technologies*), e os impactos na economia e na sociedade. Paralelamente, abordar os aspectos filosóficos da IA – entre outros os conceitos de agente, agente moral/status moral, autonomia, visão antropocêntrica e instrumental da linguagem, produção de significado, cognição, relação de co-autoria entre humanos e modelos de IA generativa -, e as barreiras para atingir a cognição humana (causalidade, raciocínio contrafactual, intencionalidade, analogia). Investigar os caminhos de mitigação dos potenciais danos à sociedade por meio de diretrizes de governança (autorregulamentação) e regulamentação pelo poder público.

2. Detalhamento da ementa

1a. semana

Apresentação da disciplina e dos alunos, metodologia e processo de avaliação de desempenho.

Introdução à disciplina.

2a. semana

Histórico e conceituação da inteligência artificial.

Fundamentos básicos da IA (alinhar conhecimento).

3a. semana

Tecnologia de propósito geral: origem e conteúdo do conceito. Histórico das GPT, ilustrando seus impactos transformadores na economia e na sociedade.

4a. semana

Conceitos de agente e agente moral com base nos autores de referência.

5a. semana

Aspectos filosóficos da inteligência artificial.

6a. semana

Aspectos filosóficos da inteligência artificial.

7a. semana

Aspectos filosóficos da inteligência artificial.

8a. semana

Barreiras para atingir a cognição humana: causalidade vs correlação.

9a. semana

Debate de textos recomendados.

10a. semana

Barreiras para atingir a cognição humana: raciocínio contrafactual, intensionalidade, analogia.

11a. semana

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

12a. Semana

Governança de IA: análise de modelos propostos e/ou adotados por organizações.

13a. semana

Regulamentação: Europa, processo e principais documentos.

14a. semana

Regulamentação: EUA, processo e principais documentos.

15a. semana

Regulamentação: China, processo e principais documentos.

16a. semana

Regulamentação: Brasil, processo e principais documentos.

17a. semana

Revisão geral e encerramento.

3. Bibliografia

3.1. Básica

BLACKMAN, Reid. Ethical Machines. Boston, Massachusetts: 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.

COECKELBERGH, Mark. The Political Philosophy of AI. USA: Polity Press, 2022.

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, Massachusetts: The MIT Press, 2012.

KAUFMAN, Dora. A inteligência artificial irá suplantar 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.

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

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

WALLACH, Wendel; Allen, Colin. *Moral Machines: Teaching Robots Right from Wrong*. UK: Oxford University Press, 2010.

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.