

# COMENTÁRIO A “POR TRÁS DA INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE DAS BASES EPISTEMOLÓGICAS DO APRENDIZADO DE MÁQUINA

*Garibaldi Sarmiento*

Professor Associado do Departamento de Filosofia da Universidade Federal da Paraíba (UFPB), João Pessoa, PB – Brasil.

 <https://orcid.org/0000-0002-9242-3945> |  [garibaldi.sarmiento@academico.ufpb.br](mailto:garibaldi.sarmiento@academico.ufpb.br)

Referência do artigo comentado: ARÃO, Cristian. Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina. **Trans/form/ação**: Revista de Filosofia da Unesp, Marília, v. 47, n. 3, e02400163, 2024. Disponível em: <https://revistas.marilia.unesp.br/index.php/transformacao/article/view/15196>.

Recebido: 10/05/2024 | Aprovado: 20/05/2024 | Publicado: 05/07/2024

 <https://doi.org/10.1590/0101-3173.2024.v47.n3.e02400221>



This is an article published in open access under a Creative Commons license

## COMENTÁRIO A “POR TRÁS DA INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE DAS BASES EPISTEMOLÓGICAS DO APRENDIZADO DE MÁQUINA”

Garibaldi Sarmento<sup>1</sup>

Referência do artigo comentado: ARÃO, Cristian. Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina. **Trans/form/ação**: Revista de Filosofia da Unesp, Marília, v. 47, n. 3, e02400163, 2024. Disponível em: <https://revistas.marilia.unesp.br/index.php/transformacao/article/view/15196>.

Cristian Arão, em seu artigo intitulado “Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina”, publicado nesta revista, afirma que “[...] será *demonstrado* como a *indução* e a *matematização* funcionam como base epistemológica da inteligência artificial e como algumas das *limitações* dessa tecnologia podem ser explicadas através das debilidades dos *métodos* que a sustentam” (Arão, 2024, p. 01, *itálicos meus*). Não obstante, o autor *não* está, *estricto senso*, apresentando uma *demonstração*, no referido artigo, pois deveríamos, para esse propósito, considerar uma análise tanto *lógico-matemática* quanto do ponto de vista de uma *refutação* metodológica, no âmbito da ciência da computação, conforme foi enfatizado em Sarmento (2004). Em princípio, as “debilidades” assinaladas são passíveis de correção, se admitirmos *a priori* a construção de um sistema maximal consistente como base de conhecimento para o modelo de *learning machines*.

Em um artigo publicado na *Revista Perspectiva Filosófica*, em 2004, intitulado “Inteligência artificial, indecidibilidade e transcomputabilidade”, eu propus uma abordagem crítica acerca dos fundamentos lógico-epistêmicos da Inteligência Artificial, na qual alguns aspectos de ordem *formal* e *técnica* – os quais fundamentam e demarcam o campo de investigações na área de Inteligência Artificial – foram tratados.

---

<sup>1</sup> Professor Associado do Departamento de Filosofia da Universidade Federal da Paraíba (UFPB), João Pessoa, PB – Brasil. ORCID: <https://orcid.org/0000-0002-9242-3945>. E-mail: [garibaldi.sarmento@academico.ufpb.br](mailto:garibaldi.sarmento@academico.ufpb.br).

Gostaria de observar inicialmente que, no âmbito das ciências empíricas, métodos de análise estatístico-probabilística são utilizados frequentemente na investigação de fenômenos do mundo físico. Não podemos, simplesmente, prescindir de uma lógica indutivo-probabilística, na prática científica. E, certamente, não devemos ignorar algumas de suas limitações formais, quando a comparamos com as lógicas dedutivas, como, p. ex., a lógica matemática clássica. Vale lembrar que, no caso da primeira, a conclusão (de um argumento *indutivo*) pode, eventualmente, ser *falsa* e suas premissas verdadeiras, o que *não* é admissível, no caso da segunda, pois, a falsidade da conclusão *é* incompatível com a veracidade das premissas. No caso de uma lógica indutivo-probabilística, no entanto, podemos estabelecer uma medida de probabilidade de *certeza* da conclusão associada à inferência indutiva.<sup>2</sup>

Outro ponto a ser reconhecido, ao contrário da linha argumentativa do autor supracitado, refere-se ao fato de que os *modelos matemáticos* são imprescindíveis para as *teorias físicas*; são esses modelos que permitem estabelecer uma interpretação empírica dos enunciados teóricos e, conseqüentemente, *validar* (ou não) as teorias físicas. Para um esclarecimento sobre a relevância epistêmica da abordagem modelo-teorética, na física atual, veja-se, p. ex., Bunge (1974).

Sem propor, neste ensaio, uma discussão de caráter técnico e formal, apresento, em linhas gerais, alguns pontos direcionados à questão da utilização de uma *lógica indutivo-probabilística* para os “Large Language Models” (LLM’s) construídos com base nas “redes neurais” que possibilitam, por sua vez, o processo de aprendizagem de máquinas.

Em primeiro lugar, algumas das áreas da ciência da computação diretamente relacionadas com os desenvolvimentos das LLM’s são a linguística computacional, a neurociência computacional e a computação quântica e, em especial, a teoria algorítmica da informação. Em ambas as áreas destacadas, métodos algorítmicos de *inferência estatística* são amplamente usados.

Uma das subáreas da teoria algorítmica da informação é denominada *complexidade* de Kolmogorov, que trata de conceitos essenciais para os métodos estatístico-probabilísticos de inferência; esse campo de estudo enseja um tratamento rigoroso do conceito de *aleatoriedade*, recorrendo-se a métodos algorítmicos, como, p. ex., o algoritmo de “probabilidade de Solomonoff”. Outro método específico para sistemas de aprendizado de máquinas, via redes neurais, é o emprego da chamada *Lógica Bayesiana*<sup>3</sup>. De um modo geral, há várias técnicas metodológicas algorítmicas que permitem um tratamento computacional rigoroso (e passível de correção) de abordagens probabilístico-estatísticas aplicáveis no desenvolvimento dos sistemas de IA.

<sup>2</sup> Ver Costa (1993), para uma abordagem filosófica e um tratamento axiomático.

<sup>3</sup> Consultar a entrada da *Stanford Encyclopedia of Philosophy*, para uma exposição introdutória acerca do Teorema de Bayes e da epistemologia bayesiana.

Sob uma perspectiva filosófica, a teoria do conhecimento compatível com o desenvolvimento desses sistemas de IA é a abordagem epistêmica *reducionista* proposta por Lehrer (1990), na qual o conceito de conhecimento se reduz, estritamente, ao conceito de informação. Lehrer admite que a explicação do *conhecimento*<sup>4</sup>, na acepção de *informação*, pressupõe que

1. *P* aceita a informação que *s*;
2. A informação que *s* seja correta;
3. A aceitação da informação que *s* seja justificada.

em que “*P*” e “*s*” são variáveis, “*P*” designa qualquer indivíduo e “*s*” denota qualquer informação que possa ser obtida. Dessa forma, “*P* sabe que *s*” expressa o “*reconhecimento da correção da informação*”. A justificativa (ou a comprovação) da informação é o aspecto central da análise epistêmica, haja vista que esse procedimento deveria ser, em princípio, entendido como uma “medida” entre uma “forte razoabilidade e *não* como certeza completa”, como defende Lehrer. Isso significa que a justificação, do ponto de vista da lógica *indutivo-probabilística*, está delimitada entre a razoabilidade e a certeza completa (o conceito de “certeza” significa aqui uma *forte* evidência de caráter comprobatório, dadas as condições *suficientes* em um certo tempo *t*, para a *confirmação* da correção da informação). Em termos probabilísticos, quanto mais próximo de 1 estiver a correção da informação, maior o grau de certeza de sua confirmação.

A teoria *algorítmica da informação* (subárea da teoria da informação), a qual se encontra na interseção da teoria da computabilidade de Turing e da teoria da informação clássica de Shannon, poderia servir de *fundamento* computacional para a abordagem advogada por Lehrer, no que tange ao *conceito* de informação e, conseqüentemente, justificar que as LLM’s processam (e geram) conhecimento sobre o “mundo físico” de modo análogo ao ser humano.

A questão filosófica principal, nesse caso, se “desloca” para o nível de *autonomia* que os sistemas cibernéticos de IA poderiam ter para *agir* de modo independente da intervenção humana. Trata-se de uma questão que, sem dúvida, requer uma reflexão mais aprofundada.

Até o presente momento, no entanto, esses sistemas cibernéticos *não* exibem efetivamente uma capacidade de autocorreção e tomada de decisões, de maneira interativa e autônoma.

Aqueles que defendem uma abordagem dessa espécie, i.e., que os sistemas cibernéticos de IA podem gerar conhecimento, não podem ignorar que a IA estará sempre *limitada* – com

---

<sup>4</sup> Uma abordagem fundamentada na teoria da *coerência* para a justificação do conhecimento.

respeito à sua capacidade cognitiva – em virtude do *teorema de incompletude* de Chaitin (uma versão do metateorema de *incompletude da Aritmética* de Gödel)<sup>5</sup>. E, por conseguinte, a IA não pode ultrapassar a capacidade cognitiva do ser humano.

## REFERÊNCIAS

- ARÃO, C. Por trás da inteligência artificial: uma análise das bases epistemológicas do aprendizado de máquina. **Trans/form/ação**: Revista de Filosofia da Unesp, Marília, v. 47, n. 2, e02400163, 2024. Disponível em: <https://revistas.marilia.unesp.br/index.php/transformacao/article/view/15196>
- BUNGE, M. **Teoria e Realidade**. São Paulo: Perspectiva, 1974.
- COSTA, N. C. A. da. **Lógica indutiva e probabilidade**. São Paulo: EDUSP – Hucitec, 1993.
- LEHRER, K. **Theory of Knowledge**. New York. Westview, 1990.
- SARMENTO, G. **Estudo das consequências epistemológicas do teorema de Gödel e do problema da palavra para semigrupos em inteligência artificial**. 1995. Dissertação (Mestrado) – UFSCar, São Carlos, SP, 1995.
- SARMENTO, G. Inteligência artificial, indecidibilidade e transcomputabilidade. **Perspectiva Filosófica**, v. I, n. 21, p. 179-209, 2004.

---

<sup>5</sup> Veja Sarmento (1994) para uma discussão lógico-epistêmica sobre o tópico.