Uma Breve Introdução à Inteligência Artificial na Física

LF EDITORIAL

Conselho Editorial da Editora Livraria da Física

Amílcar Pinto Martins – Universidade Aberta de Portugal

Arthur Belford Powell – Rutgers University, Newark, USA

Carlos Aldemir Farias da Silva – Universidade Federal do Pará

Emmánuel Lizcano Fernandes – UNED, Madri

Iran Abreu Mendes – Universidade Federal do Pará

José D'Assunção Barros – Universidade Federal Rural do Rio de Janeiro

Luis Radford – Universidade Laurentienne, Canadá

Manoel de Campos Almeida – Pontifícia Universidade Católica do Paraná

Maria Aparecida Viggiani Bicudo – Universidade Estadual Paulista - UNESP/Rio Claro

Maria da Conceição Xavier de Almeida – Universidade Federal do Rio Grande do Norte

Maria do Socorro de Sousa – Universidade Federal do Ceará

Maria Luisa Oliveras – Universidade de Granada, Espanha

Maria Marly de Oliveira – Universidade Federal Rural de Pernambuco

Raquel Gonçalves-Maia – Universidade de Lisboa

Teresa Vergani – Universidade Aberta de Portugal

Henrique Ferreira e Gustavo Martini Dalpian

Uma Breve Introdução à Inteligência Artificial na Física

LF EDITORIAL

2026

Copyright © 2026 os autores
1ª Edição

Direção editorial: Victor Pereira Marinho e José Roberto Marinho

Edição revisada segundo o Novo Acordo Ortográfico da Língua Portuguesa

Dados Internacionais de Catalogação na publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Ferreira, Henrique
Uma breve introdução à inteligência artificial na física / Henrique Ferreira, Gustavo Martini Dalpian. – São Paulo: LF Editorial, 2026.

Inclui bibliografia
ISBN 978-65-5563-822-6

1. Aprendizagem de máquina 2. Física 3. Inteligência artificial 4. Metodologia 5. Pesquisa científica I. Dalpian, Gustavo Martini. II. Título.

26-397481.0 CDD-530

Índices para catálogo sistemático:
1. Física 530

Maria Alice Ferreira - Bibliotecária - CRB-8/7964

Todos os direitos reservados. Nenhuma parte desta obra poderá ser reproduzida sejam quais forem os meios empregados sem a permissão da Editora.
Aos infratores aplicam-se as sanções previstas nos artigos 102, 104, 106 e 107 da Lei Nº 9.610, de 19 de fevereiro de 1998

LF Editorial
www.livrariadafisica.com.br
www.lfeditorial.com.br
(11) 2648-6666 | Loja do Instituto de Física da USP
(11) 3936-3413 | Editora

Agradecimentos

O presente livro nasceu de um convite do José Roberto Marinho, da Livraria da Física, em razão do crescente interesse pela temática de Inteligência Artificial e a lacuna em situar esse domínio do conhecimento com relação às Ciências Naturais como a Física. Agradecemos a oportunidade de trazer um pouco da nossa visão sobre o tema.

Parte do nosso interesse em estudar a área de Inteligência Artificial surgiu por conta da pesquisa científica que desenvolvemos utilizando essas ferramentas no âmbito da Física da Matéria Condensada. Estas pesquisas foram viabilizadas por projetos financiados pelas agências de fomento FAPESP, CAPES e CNPq, sendo este último propiciador do INCT de Informática de Materiais no qual atuamos.

A primeira vez que tivemos a oportunidade de traduzir essas pesquisas para a sala de aula foi em 2023 em um webminicurso oferecido pela Sociedade Brasileira de Física. Somos gratos à SBF pela organização e aos alunos pelos comentários. Este livro constitui uma tentativa de ampliar e sistematizar o conteúdo abordado naquele momento.

Agradecemos também aos nossos amigos Antonio Claudio Padilha, Lucas Tonetto e Tito Spadini pela leitura do manuscrito deste livro e pelas valiosas sugestões fornecidas.

E aos nossos familiares pelo apoio imprescindível.

Prefácio

Os métodos relacionados à Inteligência Artificial (IA) não são novos. Eles apareceram nas Ciências há muitos anos, e orbitam ao redor de áreas como a Computação, Matemática, Psicologia, Física e Engenharia.

Os físicos estão muito acostumados a tratar dados. Já nos primeiros anos dos cursos de graduação, em disciplinas como Laboratório de Física 1, coletam dados relativos a deslocamento e tempo em sistemas sob movimento retilíneo uniforme. Com estes dados em mãos é possível fazer um ajuste linear e obter, por exemplo, a relação entre o deslocamento e o tempo decorrido. Este tipo de procedimento, conhecido como Regressão Linear, é um tipo de algoritmo de Inteligência Artificial. Este exemplo muito simples serve para mostrar que algumas metodologias relacionadas à IA já estão disponíveis há muito tempo, e os físicos estão muito habituados a eles.

Nos últimos anos, entretanto, temos visto um avanço muito grande na IA por toda a sociedade e em todas as áreas da ciência. O Prêmio Nobel de Física e o Prêmio Nobel de Química do ano de 2024 foram relacionados a este tema, reconhecendo a importância dessa metodologia para essas áreas. Mas, considerando que estas metodologias já são tão antigas, qual o motivo de somente agora estarem despontando?

A resposta mais aceita está relacionada a dois fatores principais. Primeiro, somente nos últimos tempos passou-se

a ter bancos de dados curados e de qualidade em diversas áreas. Dados de qualidade são fundamentais para que se possa aplicar IA a um determinado problema, especialmente quando é utilizado o Aprendizado de Máquina e o Aprendizado de Máquina Profundo. O segundo fator está relacionado ao aprimoramento das peças de computadores: processadores (CPUs), memórias RAM e, principalmente, placas gráficas (GPUs). Computadores mais potentes e mais baratos ampliaram acesso ao recurso fundamental para se fazer IA, a computação.

Este pequeno livro tem por objetivo fazer uma brevíssima introdução à Inteligência Artificial, partindo de seus cânones e focando na área de Física. Considerando a complexidade do tema, e a velocidade com que ocorrem os avanços nos dias atuais, não iremos detalhar as discussões na profundidade que é usual em livros-texto. Sempre que possível citaremos referências onde o leitor poderá encontrar mais informações. Há ampla literatura sobre o tema, contemplando tanto abordagens didáticas quanto de pesquisa científica. É importante ressaltar que o pleno entendimento dessas metodologias demanda conhecimento profundo de programação, mas também de matemática e estatística. Para o presente livro, entretanto, entendemos que um estudante de graduação em física ou áreas correlatas conseguirá acompanhá-lo.

Esperamos que essa breve introdução permita aos leitores neófitos alçarem voos mais altos.

Sumário

1 Uma Síntese da IA **6**
- 1.1 História da IA 7
- 1.2 Mapa de Conceitos 24

2 Algoritmos de IA **26**
- 2.1 Regressão Linear 28
- 2.2 Regressão Logística 33
- 2.3 Vizinhos Mais Próximos 38
- 2.4 K-Médias 43
- 2.5 Máquina de Vetores de Suporte 48
- 2.6 Análise de Componentes Principais 55
- 2.7 Aprendizado por Inferência Bayesiana 59
- 2.8 Árvores de Decisão 65
- 2.9 Métodos de Ensemble 71
- 2.10 Busca em Grafos 80
- 2.11 Algoritmos Evolutivos 85
- 2.12 Sistema Algébrico Computacional 92
- 2.13 Redes Neurais Artificiais 97
- 2.14 Algoritmos de Lógica Formal 119

3 Perspectiva Integrada **127**
- 3.1 Física aplicada a problemas de IA 129
  - 3.1.1 Recozimento Simulado 129
  - 3.1.2 Redes de Hopfield 133

3.1.3 Máquinas de Boltzmann 138
3.1.4 Busca em Árvore por Monte Carlo . . 143
3.1.5 Modelos de Difusão 146
3.2 IA aplicada a problemas de Física 151
3.2.1 Reconhecimento de Galáxias 151
3.2.2 Potenciais Interatômicos 153
3.2.3 Sistemas Mecânicos Clássicos 157
3.2.4 Redes de Kolmogorov-Arnold 162
3.2.5 Regressão Simbólica 165
3.2.6 Redes Neurais Informadas por Física 170

4 Tópicos Complementares 180
4.1 Interpretabilidade e Explicabilidade 180
4.2 Processamento de Linguagem Natural 187
4.3 Robótica e IA Física 199
4.4 Ética da IA no Ensino e na Pesquisa 203

1. Uma Síntese da IA

"I'm sorry, Dave. I'm afraid I can't do that." - HAL 9000
Stanley Kubrick and Arthur C. Clark, 2001: A Space Odyssey.

Na literatura de ficção é comum que a Inteligência Artificial (IA) seja retratada por computadores e robôs humanoides que, invariavelmente, se revoltam contra seu próprio criador quando atingem um estado consciente ou um entendimento superior do mundo. Será este o fim dessa tecnologia? Fim não só como desfecho, como também finalidade no sentido teleológico. Afinal, o próprio nome Artefato não deixa mentir: é esperado que seja construído a um fim. A esse fim, o homem-projetista qualificou de Inteligência, mesmo que nem ele, apesar de sua experiência subjetiva e intuitiva, compreenda exatamente o que é. Logo, o controle, a revolta, o medo do desconhecido, são todas expressões humanas também projetadas, não na, mas sobre a máquina.

Apesar de apreciadores da literatura de ficção e do cinema, a IA do mundo real, atualmente, não poderia estar mais longe dessa visão antropomorfizada. O controle exercido pela IA está em limitar aquilo que vemos na internet. O problema existencial está em empregos que podem ser substituídos por IA. E o surgimento de uma vida inteligente como a que entendemos, ainda não se provou. Os problemas atuais são sociais e éticos[1], enquanto ainda não nos

[1]Floridi, L. A ética da inteligência artificial: princípios, desafios e oportunidades. PUCPRESS, (2025).

deparamos com os grandes problemas especulativos[2].

Neste capítulo, pretendemos desmistificar a IA por meio de seu desenvolvimento histórico, apresentando uma crônica de seus principais marcos e formulações conceituais, além de fornecer, ao final, um mapa de conceitos para que o leitor possa se orientar na literatura especializada.

1.1 História da IA

Na maioria dos livros-texto, o ano de 1956 é tido como a data inaugural da Inteligência Artificial como disciplina do conhecimento humano. Naquele ano, John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon organizaram uma oficina na universidade de Dartmouth, durante o verão, tendo em parte do título: "projeto de pesquisa em Inteligência Artificial". Foi na proposta que submeteram um ano antes, que este termo foi usado pela primeira vez[3].

Para entender melhor esse acontecimento precisamos retroceder pelo menos uma década. Em 1945 os primeiros computadores eletrônicos acabavam de se tornar uma realidade. A ideia de computador programável em si, era mais antiga, devido a *Máquina Analítica* idealizada em 1837 por Charles Babbage e Ada Lovelace. Porém, esse computador puramente mecânico nunca chegou a ser produzido completamente. A computação mecânica realizada no final do século XIX era analógica e não programável (por exemplo, a máquina de prever marés inventada pelo físico William Thomson, Lord Kelvin[4]). A formalização do conceito de

[2]Bostrom, N. Superintelligence: Paths, Dangers, Strategies. Oxford University Press, (2016).

[3]McCarthy, J., Minsky, M. L., Rochester, N. & Shannon, C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Mag. 27, 12 (2006).

[4]Ferrel, W. The Maxima and Minima Tide-Predicting Machine. Science ns-3, 408–410 (1884).



Figura 1.1: *Dartmouth Summer Research Project on Artificial Intelligence*, 1956. Da esquerda para a direita, Oliver Selfridge, Nathaniel Rochester, Ray Solomonoff, Marvin Minsky, Peter M. Milner, John McCarthy e Claude Shannon. (Fonte: IEEE.ORG / The Minsky Family).

computador só ocorreria em novembro de 1936, proposta pelo matemático Alan Turing[5].

Foi a II Guerra Mundial que direcionou investimentos para a construção de computadores que operavam eletricamente, eram digitais e programáveis, como o *ENIAC* e o *Colossus*. Tais esforços também influenciaram na Teoria de Controle, que avançou para ao aperfeiçoamento de artilharia. Concomitantemente, o crescimento comercial da telefonia

[5]Turing, A. M. On Computable Numbers, with an Application to the Entscheidungsproblem. Proc. Lond. Math. Soc. s2-42, 230–265 (1937).

e do rádio fomentou a criação da Teoria da Informação[6]. O polímata Nobert Wiener denominou Cibernética a união desses ramos: computação, controle e comunicação[7]. De fato, o primeiro modelo matemático de neurônio foi proposto em 1943 pelo autointitulado ciberneticista Warren McCulloch, em colaboração com Walter Pitts[8]. McCulloch organizou entre 1941 e 1960 as, posteriormente nomeadas, Conferências de Cibernética, reunindo matemáticos, engenheiros, psicólogos e neurocientistas. Em uma dessas conferências, em 1951, Shannon apresentou um mecanismo eletro-eletrônico capaz de resolver um labirinto sozinho[9], um óbvio exemplo de um artefato do que, em quatro anos, seria batizado por McCarthy de Inteligência Artificial. A escolha do nome foi proposital: McCarthy queria distinguir de um nome utilizado em um livro co-editado por ele e Shannon (*Automata Studies*), ao mesmo tempo que queria se distanciar do termo cibernética, cuja autoridade de Wiener era incontestável[10].

Nesse caldeirão de ideias, a pergunta que borbulhava era exatamente se "as máquinas poderiam pensar". Turing não hesitou: propôs um teste para verificar essa possibilidade, o hoje conhecido Teste de Turing[11]. A versão moderna do teste consiste em uma situação na qual um interrogador humano se comunica, por meio exclusivamente textual, com dois interlocutores ocultos, um humano e uma máquina. O objetivo do interrogador é determinar qual dos dois é

[6]Shannon, C. E. A Mathematical Theory of Communication. Bell Syst. Tech. J. 27, 379–423, 623–656 (1948).

[7]Wiener, N. Cibernética e Sociedade: o uso humano de seres humanos. Cultrix (1954).

[8]McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5, 115–133 (1943).

[9]Gleick, J. A Informação, Companhia das Letras (2013).

[10]Nilsson, N. J. The Quest for Artificial Intelligence. Cambridge University Press (2010).

[11]Turing, A. M. Computing Machinery and Intelligence. Mind 49, 433–460 (1950).

o humano. Se, após uma série de perguntas livres, o interrogador não conseguir distinguir de maneira confiável a máquina do participante humano, diz-se que a máquina foi bem-sucedida no teste. Com isso, Turing propôs substituir a questão abstrata "as máquinas podem pensar?" por um critério operacional baseado no desempenho linguístico e comportamental. Como o interrogador pode formular perguntas de qualquer natureza, o teste exige da máquina respostas flexíveis e adaptativas, comparáveis às de um ser humano.

Enquanto a Cibernética dava ênfase a sistemas analógicos adaptativos baseados em retroalimentação, McCarthy preferia a abordagem simbólica permitida pelas emergentes linguagens de programação. No *workshop* de Dartmouth, outros participantes compartilhavam dessa opinião, entre eles, Marvin Minsky, Herbert Simon e Allen Newell. É creditado a estes dois últimos o primeiro código de IA, o *Logic Theorist*, escrito em 1956 e capaz de provar teoremas matemáticos de forma automática[12]. Essa abordagem é conhecida como raciocínio simbólico, fortemente ligada a lógica formal, constituindo-se em uma das abordagens possíveis dentro do campo da IA.

Apesar da concordância quanto à abordagem, Simon e Newell estavam mais interessados em aspectos psicológicos do raciocínio, enquanto McCarthy e Minsky exploravam a abordagem simbólica sem fazer referência específica ao comportamento humano.

Este tipo de distinção entre fidelidade ao comportamento humano e ao raciocínio correto carateriza quatro dimensões de como a IA pode ser entendida[13]. Para alguns pesquisadores o foco é no comportamento que pode ser visto externa-

[12]Newell, A. & Simon, H. The Logic Theory Machine — A Complex Information Processing System. IEEE Trans. Inf. Theory 2, 61–79 (1956).

[13]Russell, S. & Norvig, P. Artificial Intelligence. 4th ed. Pearson (2020).

mente: o objetivo da máquina é emular os resultados de um ser humano naquela tarefa. Para outros, o foco é em obter os melhores resultados possíveis em um cenário que exige raciocínio. Neste caso, a máquina deve "pensar racionalmente", como se a tarefa pudesse ser reescrita na forma de um problema de otimização. Para alguns outros pesquisadores, entretanto, construir artefatos de IA pode significar simular o pensamento humano, no sentido psicológico do termo. Isso está mais próximo da Psicologia Cognitiva que assume um modelo computacional para explicar os processos cognitivos humanos, dando origem a terminologia Inteligência Computacional[14]. Por fim, os artefatos de IA também podem "agir racionalmente", no sentido do comportamento externo observado. Isso significa que o processo interno não é, necessariamente, matematicamente ótimo ou com uma lógica inferencial explícita; contudo, o comportamento externo observado apresenta padrões racionais que não correspondem aos padrões humanos de comportamento. Por exemplo, o método pelo qual formigas e outros insetos sociais encontram rotas sub-ótimas representa um comportamento racional, apesar de formigas não raciocinarem individualmente sobre a rota que "devem escolher" (em IA isso é chamado de Inteligência de Enxame[15]).

Entre 1964 e 1966, Joseph Weizenbaum utilizou a abordagem simbólica para criar o *ELIZA*, um sistema que imitava o comportamento humano em uma conversa e que se configurou como o primeiro *chatbot* construído[16]. Era uma primeira tentativa de resolver o Teste de Turing emulando inteligência.

[14]Feigenbaum, E. A. Some Challenges and Grand Challenges for Computational Intelligence. J. ACM 50, 32–40 (2003).

[15]Beni, G. & Wang, J. Swarm Intelligence in Cellular Robotic Systems. 703–712. Springer (1993).

[16]Weizenbaum, J. ELIZA — a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 36–45 (1966).

Por outro lado, os ciberneticistas continuavam a avançar por meio da eletrônica analógica. Em 1958, Frank Rosenblatt havia implementado a primeira rede neural artificial baseado nos modelos de McCulloch e Pitts[17]. O *Perceptron* foi primeiramente testado na forma de código em um *IBM 704*, e depois construído como um emaranhado de fios e potenciômetros (o *MARK I Perceptron*). Em ambos os casos, o sistema era capaz de aprender como classificar baseando-se em exemplos de treinamento. A retroalimentação demonstrava sua força em criar sistemas inteligentes.

As formas como o treinamento poderia ser executado eram variadas, e logo outras opções foram desenvolvidas, como o *Adaptive Linear Neuron* (*ADALINE*) proposto em 1960 por Bernard Widrow[18] e seu aluno de doutorado Marcian Hoff. O *ADALINE* introduziu o uso sistemático da regra do erro mínimo quadrático para ajustar os pesos sinápticos do neurônio artificial, o que é amplamente utilizado até os dias de hoje.

O engenheiro elétrico Arthur Samuel tinha suas próprias ideias do que deveria ser o Aprendizado de Máquina. Em 1958, inspirado na ideia de Shannon de desenvolver um sistema capaz de jogar xadrez, Samuel criou um programa que aprendia a jogar damas à medida que analisava diversas partidas, ajustando seus parâmetros com base na experiência acumulada[19]. Em vez de ser uma abordagem neural, com ajuste de pesos sinápticos como o *Perceptron* e o *ADALINE*, a proposta de Samuel criava uma representação simbólica do tabuleiro através de árvores (uma estrutura de dados

[17]Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. Psychol. Rev. 65, 386–408 (1958).

[18]Widrow, B. An Adaptive ADALINE Neuron Using Chemical Memistors. Technical Report No. 1553-2, Stanford Electronics Laboratories, Stanford University (1960).

[19]Samuel, A. L. Some Studies in Machine Learning Using the Game of Checkers. IBM J. Res. Dev. 3, 210–229 (1959).

computacionais) e tinha uma função de avaliação heurística explícita. Ou seja, a aprendizagem ajustava parâmetros dentro uma estrutura simbólica já definida.

Em 1969, Minsky e Seymour Papert publicam um livro para discutir a teoria matemática por trás dos *Perceptrons*[20]. O livro teve grande impacto na comunidade de Inteligência Artificial devido a demonstração matemática da impossibilidade dos *Perceptrons* resolverem a operação lógica denominado ou-exclusivo. Minsky era tido como uma autoridade em IA, de maneira que as pesquisas em redes neurais artificiais minguaram[21].

Esse alvorecer da IA ainda testemunharia o surgimento de aplicações comerciais dos denominados Sistemas Especialistas. Construídos em cima da abordagem simbólica, estes sistemas executavam tarefas que antes exigiriam um especialista no assunto. Sua construção era baseada em uma representação de conhecimento que pudesse ser armazenada e manipulada para gerar inferências lógicas. Os denominados engenheiros de conhecimento construíam essa representação à medida que anotavam regras e heurísticas comentadas pelos especialistas. Os dois principais exemplos foram desenvolvidos na universidade de Stanford entre 1960 e 1980: o sistema *DENDRAL* foi criado por Edward Feigenbaum e colaboradores, projetado para atuar como um especialista em química analisando espectrogramas de massa para determinar a estrutura de moléculas[22]; já o sistema *MYCIN*, desenvolvido por Edward Shortliffe, atuava no domínio das infecções bacterianas, diagnosticando e sugerindo tratamentos[23].

---

[20] Minsky, M. & Papert, S. Perceptrons. MIT Press (1969).

[21] Domingos, P. O Algoritmo Mestre. Novatec Editora (2017).

[22] Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A. & Lederberg, J. DENDRAL: A Case Study of the First Expert System for Scientific Hypothesis Formation. Artif. Intell. 61, 209–261 (1993).

[23] Shortliffe, E. H. Computer-Based Medical Consultations: MYCIN.

Também na robótica as abordagens de IA eram majoritariamente simbólicas. Entre 1966 e 1972, financiados pelo *DARPA*[24], Charles Rosen, Nils Nilsson e Peter Hart desenvolveram o primeiro robô móvel de propósito geral capaz de raciocinar sobre suas próprias ações, o *Shakey the Robot*[25]. Durante o processo, Nilsson e colaboradores criaram o algoritmo A* (A estrela) para resolver problemas de planejamento de caminhos e busca em grafos, permitindo que o robô avaliasse diferentes rotas possíveis, estimasse o custo até o objetivo e escolhesse o caminho mais eficiente. O A* é um dos algoritmos de busca mais importantes da Inteligência Artificial, usado até hoje em robótica, jogos digitais, GPS e planejamento de rotas.

Em 1969, Marvin Minsky foi agraciado com o Prêmio Turing, "por seu papel central na criação, no desenvolvimento, na promoção e no avanço do campo da Inteligência Artificial". A premiação havia sido criada em 1966 pela *Association for Computing Machinery*, em homenagem a Alan Turing, e é atualmente considerada equivalente em prestigio a um prêmio Nobel da computação. Em 1971, John McCarthy era quem recebia o prêmio. Na página da premiação, sua láurea é creditada ao seu papel fundamental na criação, consolidação e promoção do campo da Inteligência Artificial, incluindo a introdução do próprio termo, o desenvolvimento da linguagem de programação LISP e a liderança de iniciativas pioneiras que estabeleceram as bases científicas e tecnológicas da área.

As comemorações, contudo, prenunciavam tempos difíceis. De 1974 a 1980, um inverno abateu o campo. Em

---

Elsevier Science Ltd (1976).

[24] *Defense Advanced Research Projects Agency*, agência de pesquisa e desenvolvimento ligada ao Departamento de Defesa dos Estados Unidos.

[25] Nilsson, N. J. (ed.) Shakey the Robot. Technical Note 323, Artificial Intelligence Center, Computer Science and Technology Division, SRI International (1984).

1973, o matemático James Lighthill apresentou um relatório para o *Science Research Council*, um órgão do governo britânico responsável por financiar pesquisas[26]. O relatório foi amplamente desfavorável ao campo de IA, assinalando que os métodos utilizados enfrentavam explosão combinatória com o tamanho do problema tratado, os sistemas ainda eram fortemente dependentes de heurísticas humanas inseridas manualmente caso a caso, e, principalmente, que muitos dos exemplos criados eram *toy models* não havendo uma evidência concreta de progresso rumo a uma inteligência geral comparável à humana.

Outro inverno viria entre 1987 e 1993, dessa vez devido ao colapso do mercado da principal aplicação comercial de IA. Os Sistemas Especialistas requeriam computadores dedicados denominados *LISP machines*, que eram programados para resolver problemas específicos das empresas. Porém, o desenvolvimento dessas aplicações era caro, os sistemas eram frágeis fora do domínio bastante restrito da aplicação, e o retorno financeiro real era muito baixo.

Estes dois invernos decretaram o fim do que ficou conhecido como *GOFAI*, *Good Old-Fashioned AI*, a linha dominante de pesquisa cuja principal abordagem era simbólica.

Entrincheirados desde muito antes, os conexionistas, herdeiros da linha ciberneticista, continuaram trabalhando com redes neurais artificiais durante esse tempo. A questão levantada por Minsky e Papert havia sido solucionada com redes envolvendo maior quantidade de camadas de neurônios artificiais. Porém, outro problema impedia o avanço: como atualizar os pesos sinápticos de todas essas camadas da rede durante o aprendizado? A solução tornou-se amplamente conhecida apenas em 1986, com o algoritmo de *backpropagation* proposto por David Rumelhart, Geoffrey Hinton e Ronald

[26] Lighthill, J. Artificial intelligence: a general survey. Science Research Council, Artificial Intelligence: A Paper Symposium. SRC (1973)

Williams[27]. Como muitas vezes na história da ciência, a ideia já havia aparecido antes (em sua versão moderna, pelo menos duas vezes: na dissertação de mestrado do matemático Seppo Linnainmaa[28] e na tese de doutorado do cientista social Paul Werbos[29]), mas foi o trabalho de Rumelhart que popularizou a técnica.

Em 1980, Kunihiko Fukushima introduziu o *Neocognitron*, uma rede multicamada para reconhecimento de imagens que empregava uma hierarquia de neurônios específicos[30]. A ideia inspirou Yann LeCunn, que entre 1988 e 1998, criou várias versões dessa rede neural usando o algoritmo de *back-propagation*. Eram as primeiras redes neurais convolucionais. Logo, novas arquiteturas de redes, com neurônios diferentes, ligações fazendo caminhos circulares, cruzados, ou mesmo ligações só para frente, dominaram as experiências dessa tecnociência.

Paralelamente outra abordagem tornava-se popular, as Máquinas de Vetor Suporte (do inglês, *Support Vector Machines*, SVM), invetadas principalmente pelos matemáticos russos Vladimir Vapnik e Alexey Chervonenkis[31]. Partindo de uma abordagem matemática abstrata que nada tinha a ver com neurônios artificiais, as SVMs conseguiam alcançar performance superior as Redes Neurais Artificias em determinadas tarefas.

---

[27] Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. Nature 323, 533–536 (1986).

[28] Linnainmaa, S. The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors. Master's Thesis, University of Helsinki, pp. 6–7 (1970).

[29] Werbos, P. J. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. PhD Thesis, Harvard University (1974).

[30] Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybern. 36, 193–202 (1980).

[31] Cortes, C. & Vapnik, V. Support-vector networks. Mach. Learn. 20, 273-297 (1995).