

# Inteligența Artificială și Premiul Nobel în Fizică (2024)

Catalin VRABIE / 24.02.2025

Recent (octombrie 2024), Premiul Nobel pentru Fizică a fost acordat lui John Hopfield de la Universitatea *Princeton* din Statele Unite și lui Geoffrey Hinton de la Universitatea din *Toronto*, Canada, pentru contribuțiile fundamentale în cercetarea și dezvoltarea rețelelor neuronale (NN) - modele de bază ce alimentează sistemele de inteligență artificială (AI), *machine learning* (ML) și toate celelalte tehnologii asociate [1].

Dacă John Hopfield este fără îndoială fizician, Geoffrey Hinton însă este specialist în știința calculatoarelor (altfel spus, informatică) iar acest din urmă fapt a provocat ceva agitație în comunitățile științifice de pe mapamond – absolut inutil, aș spune eu.

Prin acordarea Premiului, Comitetul Nobel a făcut o declarație foarte puternică și anume că tehnologiile *machine learning* și inteligența artificială în general, datorează foarte mult fizicii. De fapt, fizica le-a inspirat, iar ceea ce numim astăzi informatică și AI își găsesc începuturile în fizică. M-am lansat în redactarea acestui articol tocmai pentru a explica puțin istoria acestei colaborări științifice<sup>1</sup> și a vă permite să reflectați asupra actualelor granițe ale disciplinelor contemporane.



Ceremonia de decernare a Premiilor Nobel 2024 (sursa: Nobel Prize YouTube Channel)

Nu există un Premiu Nobel pentru informatică<sup>2</sup>; totuși în ultimii ani, descoperirile importante din chimie sau astronomie, au avut la bază algoritmi programați de specialiști informaticieni. În fapt, chiar mesajul oficial ce anunța cel mai recent Premiu Nobel în Fizică amintește motivul acordării *for*

<sup>1</sup> Recomand cititorilor interesați de mai multe detalii, parcurgerea volumului „AI de la idee la implementare. Traseul sinuos al Inteligenței Artificiale către maturitate” în care este prezentată o istorie mai largă a domeniului inteligenței artificiale [5].

<sup>2</sup> Există însă un echivalent al acestuia – *ACM A.M. Turing Award*, câștigat în 2019 de echipa formată din Yoshua Bengio, Geoffrey Hinton și Yann LeCun pentru *Major Breakthroughs in Artificial Intelligence* [12].

*foundational discoveries and inventions that enable machine learning with artificial neural networks*<sup>3</sup> [1].

## Cine sunt laureații?

John J. Hopfield

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



© Nobel Prize Outreach. Photo: Nanaka Adachi

Geoffrey Hinton

“for foundational discoveries and inventions that enable machine learning with artificial neural networks”



© Nobel Prize Outreach. Photo: Clément Morin

Sursa: [nobelprize.org](https://nobelprize.org)

John Hopfield este un fizician care, la începutul carierei sale, a descoperit polaritonii [2]; dar a fost premiat de Nobel pentru modelul Hopfield. Acesta este un model al memoriei asociative, altfel spus modul în care se poate defini un sistem ce funcționează ca o memorie putând recunoaște tipare cunoscute chiar și dintr-o reprezentare imperfectă a acestora (ca atunci când reconstruim din amintiri drumul spre hotelul unde suntem cazați, chiar dacă între timp condițiile meteo s-au schimbat).

Geoffrey Hinton, pe de altă parte, este un specialist în neuroștiințe și un informatician extrem de influent, un veritabil pionier al inteligenței artificiale, care a realizat (direct sau indirect) multe dintre primele lucrări din acest domeniu. Recunoașterea Nobel a venit pentru „mașinile Boltzmann”, modele foarte asemănătoare cu cele ale lui Hopfield, dar concepute pentru a realiza ceea ce în *machine learning* se numește *unsupervised learning* – sisteme capabile să învețe din date și apoi să genereze noi exemple compatibile cu datele din care au învățat. Elementele de bază sunt aceleași, ba chiar și numele Boltzmann sugerează că ambele modele folosesc principii fundamentale fizicii [3, 4].

Totuși, au existat numeroase discuții, dezbateri și controverse online despre faptul că această cercetare ar aparține (sau nu) cu adevărat fizicii. Unii susțin că teoria este pur informatică (inspirată de neuroștiințe); dar, dacă ne gândim bine, forma cea mai pură a informaticii este adesea de natură matematică, ceea ce o face să semene foarte mult cu matematica însăși. Există însă și informatică aplicată, care se suprapune cu multe alte domenii. Așadar, este oare relevantă și utilă o asemenea dezbatere ori încercarea de „a închide” teoriile în anumite discipline?

În opinia mea, nu. Premiul Nobel în Fizică din 2024 a fost acordat, în esență, pentru algoritmi informatici inspirați din teoriile fizicii. Aceștia au condus la noi descoperiri, care la rândul-le au generat altele și altele – până când, astăzi, la începutul anului 2025, *Generative AI* (GenAI) ne ajută să creăm lumi fantastice stând în sufragerie, în fața unui telefon *smart* [5].

Revenind la controversă: este adevărat că nu a fost descoperită „cea mai cea particulă din univers” (aluzie la scrierile lui Douglas Adams<sup>4</sup>) și nici măcar ceva similar ei, dar, esența cercetării premiate de Comitetul Nobel este, fără îndoială, fizica.

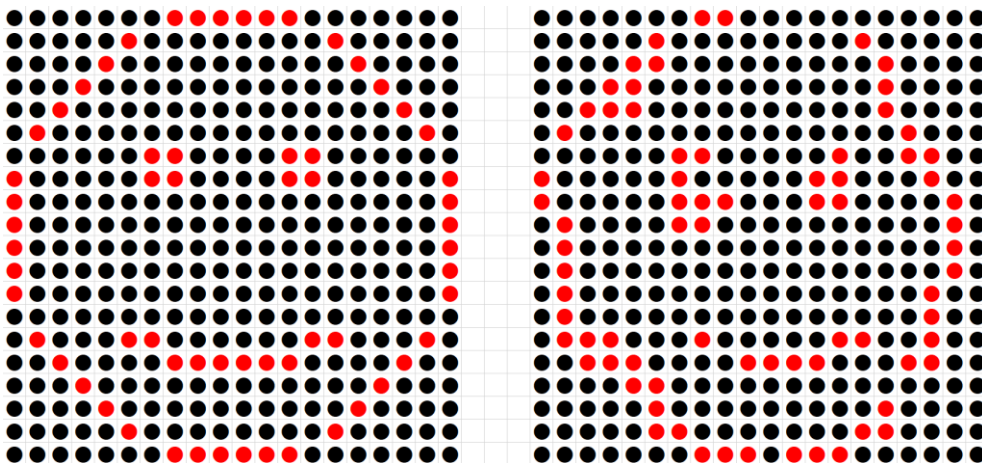
<sup>3</sup> „Descoperiri și invenții fundamentale care permit învățarea automată cu rețele neuronale artificiale”.

<sup>4</sup> Fac aici referire la volumul *The Hitchhiker’s Guide to the Galaxy* al acestui autor, volum pe care îl recomand spre citire.

Există un minunat articol, publicat în martie 2014 de John Hopfield în *Annual Review of Condensed Matter Physics*, intitulat *Whatever Happened to Solid-State Physics?* [6] ce trece prin întreaga istorie a teoriei Hopfield (și a autorului ei) care merită citit. Rezumând în câteva cuvinte, ceea ce este special la Hopfield este faptul că primii săi pași în domeniul fizicii au fost făcuți prin intermediul bateriilor, bobinelor, construirii de radiouri etc. – elemente simple, dar fundamentale. Întrucâtva, cred că fizicienii au tendința de a reduce lucrurile la cele mai simple componente posibile – fie că este vorba despre particule fundamentale ale materiei sau despre ecuații elegante care descriu diverse fenomene. De multe ori însă, cele mai interesante aspecte nu țin doar de particulele în sine, ci de interacțiunile dintre ele. Un comportament simplu la nivel de particule, când este văzut la scară mare (așa cum se întâmplă într-o rețea neuronală), dă naștere unor comportamente incredibil de fascinante, importante și, mai presus de orice, funcționale și poate, într-o oarecare măsură, replicabile.

### Ce este o rețea Hopfield?

Ei bine, aici apar cu adevărat legăturile dintre fizică și informatică; pentru o mai bună înțelegere voi descrie pe scurt ceva denumit modelul Ising [7]. Acest model a fost dezvoltat inițial pentru a explica magnetismul – pentru a înțelege de ce unele materiale sunt magnetice, iar altele nu. Magnetismul, în esență, se bazează pe o caracteristică a electronilor, dar cum nu intenționez să dau mai multe explicații în această direcție (fără a uita să invit totuși cititorii curioși să investigheze fenomenul pe internet sau în biblioteci) continui spunând că această caracteristică se numește *spin* (care poate fi orientat în sus sau în jos). Dacă reprezentăm acest lucru folosind jetoane colorate cu negru pe o față și roșu pe cealaltă pe care le așezăm pe o masă, putem reduce un sistem incredibil de complex – cu miliarde și miliarde de electroni, la un model mult mai simplu. La prima vedere, un sistem cu doar două stări, spin sus sau spin jos (roșu / negru), pare trivial. Dar când luăm în calcul interacțiunile dintre aceste stări, lucrurile încep să devină cu adevărat interesante.



Un exemplu de aranjare a jetoanelor noastre. În stânga avem un *smiley face* (starea fundamentală); în dreapta o versiune distorsionată a lui

Cum se leagă cercetarea pentru care Hopfield a primit Premiul Nobel de modelul Ising și ce legătură este între magnetism și rețelele neuronale? Ei bine, ceea ce a făcut Hopfield a fost să dezvolte, probabil, cea mai simplă versiune posibilă a unei rețele neuronale. El a fost interesat de modul în care o rețea neuronală, precum cea din creierul nostru, poate funcționa – prin contrast de cum funcționează cele ale computerelor care se bazează pe locații specifice și *pointer*-i [8].

Dacă un tipar (precum cel din imaginea de mai sus, stânga) ar fi corupt sau distorsionat (așa cum apare în dreapta), ar putea totuși să fie recunoscut de model? Hopfield a creat un sistem în care un

anumit tipar este definit ca „starea fundamentală<sup>5</sup>” – acesta fiind ceea ce dorim să reținem, ceea ce trebuie să fie memorat. El a stabilit valoarea conexiunilor sau interacțiunilor dintre neuroni astfel încât acestea să corespundă stării fundamentale. Dacă alimentăm sistemul cu o versiune distorsionată a tiparului original și ajustăm valorile legăturilor dintre neuroni astfel încât energia de care avem nevoie pentru a ajunge la starea fundamentală să scadă, păstrăm configurația. Dacă energia crește, o respingem. În cele din urmă, sistemul va evolua spre tiparul original, deoarece acesta a fost cel setat ca fiind starea fundamentală. În acest fel rețeaua se ajustează singură pentru a reduce energia și pentru a recupera modelul original.

O întrebare naturală care se ridică este: câte tipare pot fi codificate într-un model Hopfield? Un rezultat important în acest domeniu a fost calcularea capacității, adică a numărului de tipare pe care un model de o anumită dimensiune le poate încorpora. Instrumentele matematice necesare pentru a rezolva această problemă au fost dezvoltate de Giorgio Parisi, care a câștigat Premiul Nobel pentru Fizică în 2021 [9]. Există așadar o conexiune directă foarte clară între eforturile lui Hopfield și fizică – aceasta fiind știința care a oferit instrumentele necesare rezolvării problemei Hopfield.

Între timp, Geoffrey Hinton a făcut progrese enorme în domeniul inteligenței artificiale ducând mai departe dezvoltarea rețelelor neuronale prin crearea „mașinilor Boltzmann”. Aceste modele, destul de dificil de utilizat în forma lor inițială, au fost restructurate într-o variantă mai eficientă numită „mașini Boltzmann restricționate”, care seamănă tot mai mult cu o rețea neuronală artificială – prin adăugarea de straturi suplimentare, rețelele create evoluează treptat și devin din ce în ce mai eficiente. Mașinile Boltzmann [3] sunt o extensie a rețelelor Hopfield, fiind în esență, modele generative, similare conceptual cu ceea ce vedem astăzi, când generăm imagini sau text cu *ChatGPT*. La acea vreme însă, erau generate mostre noi de imagini foarte simple sau mici seturi noi de date.

În justificarea Premiului Nobel se face trimitere la una dintre contribuțiile majore ale lui Geoffrey Hinton și anume popularizarea utilizării algoritmului de *backpropagation* – un algoritm care permite rețelelor neuronale să se îmbunătățească în timp, permițând modelului să își corecteze greșelile pentru a obține rezultate mai bune în iterațiile viitoare [10]. Acest algoritm este incredibil de răspândit – practic, fiecare rețea neuronală de astăzi folosește *backpropagation*. Indiferent dacă vorbim despre *reinforcement learning*, *deep learning* sau *LLMs*<sup>6</sup>, în cele din urmă, toate folosesc o formă de *backpropagation* pentru a-și ajusta parametrii și a deveni mai performante.

Una dintre cele mai citate lucrări ale lui Hopfield, și anume: *Neural Networks and Physical Systems with Emergent Collective Computational Abilities* [11], publicată în ianuarie 1982, face clară încă din titlu, legătura dintre sistemele fizice, fizică și rețele neuronale. Desigur, există și câteva limitări, dar ideea de a reprezenta o rețea neuronală ca un sistem binar cu interacțiuni între elementele sale a fost preluată și dezvoltată apoi de Hinton și de mulți alți oameni de știință. Această bază a dus, în cele din urmă, la o explozie de cercetări și aplicații, motiv pentru care rețelele neuronale artificiale au devenit astăzi omniprezente.

Ar fi nerealist să presupunem că am fi avut același număr de descoperiri revoluționare în orice alt domeniu dacă nu ar fi fost susținute, la un moment dat, de știința calculatoarelor. Cercetarea interdisciplinară este locul în care apar cele mai mari descoperiri. Fie că vorbim despre chimie (Premiul Nobel pentru Chimie din 2024 a fost acordat dezvoltatorilor *AlphaFold*, în esență, o aplicație *deep learning* pentru chimie – voi detalia acest subiect într-un articol viitor), fizică, neuroștiințe sau inteligență artificială, cele mai influente lucrări sunt cele la intersecția mai multor discipline.

---

<sup>5</sup> *Lowest energy state* așa cum este cunoscut termenul în literatura de specialitate.

<sup>6</sup> *Large Language Models*

Există o mulțime de cercetări în desfășurare la interfața dintre fizică – dar cu siguranță putem spune asta despre toate ramurile științei, și domeniul inteligenței artificiale, împrumutându-se idei și concepte dintr-un domeniu în celălalt. Această colaborare este incredibil de intensă și astăzi, mai prețioasă ca niciodată.

## References

- [1] The Nobel Prize, "Nobel Prize in Physics 2024," The Nobel Prize, 2024. [Online]. Available: <https://www.nobelprize.org/prizes/physics/2024/summary/>. [Accessed 21 02 2025].
- [2] M. Litinskaya, "Polaritons," in *Encyclopedia of Condensed Matter Physics*, Elsevier, 2005, pp. 334-341.
- [3] G. Hinton, "Boltzmann Machines," 25 03 2007. [Online]. Available: <https://www.cs.toronto.edu/~hinton/csc321/readings/boltz321.pdf>. [Accessed 21 02 2025].
- [4] G. Hinton and T. Sejnowski, "Optimal perceptual inference," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Washington, 1983.
- [5] C. Vrabie, AI de la idee la implementare. Traseul sinuos al Inteligenței Artificiale către maturitate. [AI from idea to implementation. The winding path of Artificial Intelligence to maturity], Bucharest: Pro Universitaria, 2024.
- [6] J. Hopfield, "Whatever Happened to Solid State Physics?," *Annual Review of Condensed Matter Physics*, vol. 5, 2014.
- [7] Stanford University, "The Ising Model," Stanford University, 20 09 2020. [Online]. Available: <https://stanford.edu/~jeffjar/statmech/intro4.html>. [Accessed 21 02 2025].
- [8] C. Vrabie, Elemente de IT pentru Administrația Publică, București: Pro Universitaria, 2024.
- [9] The Nobel Prize, "Nobel Prize in Physics 2021," The Nobel Prize, 2021. [Online]. Available: <https://www.nobelprize.org/prizes/physics/2021/summary/>. [Accessed 21 02 2025].
- [10] D. Bergmann and C. Stryker, "What is backpropagation?," IBM, 02 06 2024. [Online]. Available: <https://www.ibm.com/think/topics/backpropagation>. [Accessed 21 02 2025].
- [11] J. Hopfield, "Neural networks and physical systems with emergent collectivecomputational abilities," in *Proc. Natl Acad. Sci. USA*, 1982.
- [12] Association for Computing Machinery (ACM), "2018 ACM A.M. Turing Award Laureates," ACM, 2018. [Online]. Available: <https://awards.acm.org/about/2018-turing>. [Accessed 02 21 2025].