



49 Zjazd Fizyków Polskich
Katowice 2025

Fizyka w AI, AI w fizyce

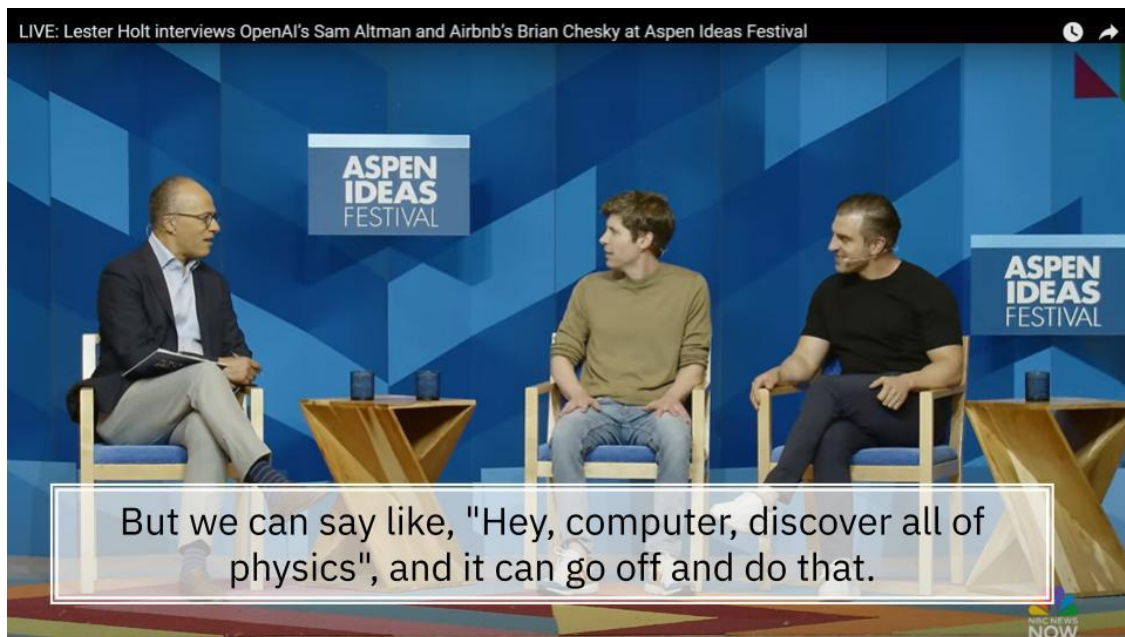
Julian Sienkiewicz

Wydział Fizyki Politechniki Warszawskiej

9 września 2025 r.



26 czerwca 2024



8 lipca 2024

Czy kiedy AI stanie się popularna wśród naukowców, czeka nas nieznane dotąd przyspieszenie rozwoju technologicznego? W jaki sposób asystenci AI już wkrótce zmienią sposób, w jaki się uczymy, pracujemy, odpoczywamy, żyjemy? **Czy sztuczna inteligencja „rozwiąże” fizykę lub matematykę?** I w jaki sposób powinniśmy rozwijać sztuczną inteligencję, żeby przyniosła nam tylko korzyści?

12 września 2024



 OpenAI

16 września 2024



A tu, proszę, przewidywania, o których mówiłem parę miesięcy temu w dyskusji m.in. z Aleksandrem Mądrym z OpenAI się sprawdziły 😊 Najnowszy model GPT-o1 ma dokładnie to, czego brakowało wcześniejszym wersjom, czyli coś na kształt pierwszego w historii sztucznego "Systemu-2" wzorowanego na procesach rozumowania znanych u ludzi (vide prace Daniela Kahnemana). Efekt jest taki, że algorytm rozwiązuje bardziej zaawansowane problemy z fizyki teoretycznej niż te, które dajemy studentom na kolokwiah, a pada dopiero na naprawdę podchwytliwych i nieortodoksyjnych problemach...

Jak by ktoś chciał się sprawdzić, można guglować, to niezbyt pomoże: podaj pierwszy nieznikający wyraz rozwinięcia wyrażenia na całkowity kąt obrotu w precesji Thomasa w ruchu elektronu na klasycznej orbicie s w atomie wodoru przy założeniu stałej prędkości i pojedynczego pełnego cyklu orbitalnego.

Wyprowadzenie poprawnego wyniku zajęło GPT-o1 13 sekund, mimo że zadanie właśnie wymyśliłem, tak, że ten 😊



W czterech rogach kwadratowego sufitu o krawędzi a znajdują się pająki, które w pewnej chwili zaczynają się poruszać w ten sposób, że każdy z nich kroczy ze stałą wartością prędkości v , utrzymując kierunek na swojego poprzednika. Po jakim torze poruszać się będą pająki i po jakim czasie nastąpi ich spotkanie? Jaką drogę przebędzie każdy z nich do momentu spotkania? Pokaż obliczenia.



Myśli dotyczące 29 sekund(y) ▾

Podsumowanie

1. **Tor ruchu:** Pająki poruszają się po spiralach logarytmicznych, zbliżając się do środka kwadratu.
2. **Czas spotkania:** Pająki spotkają się po czasie:

$$T = \frac{a}{v}$$

3. **Przebyta droga:** Każdy pająk przebywa drogę równą:

$$S = a$$

1. Tor ruchu pajaków

a) Opis ruchu

Pająki znajdują się w wierzchołkach kwadratu o boku a . Każdy pająk porusza się z prędkością v w kierunku swojego poprzednika, tworząc zamkniętą pętlę. Ze względu na symetrię problemu, trajektorie pajaków będą identyczne i skierowane spiralnie w stronę środka kwadratu.

b) Model matematyczny ruchu

Aby opisać tor ruchu pajaków, użyjemy układu współrzędnych biegunowych z początkiem w środku kwadratu. Pozycja każdego pajaka w czasie t będzie opisana przez współrzędne $(r(t), \theta(t))$, gdzie:

- $r(t)$ — odległość od środka kwadratu,
- $\theta(t)$ — kąt względem osi poziomej.

d) Wyprowadzenie równania trajektorii

Aby znaleźć zależność między r a θ , podzielimy równania:

$$\frac{dr}{dt} = -\frac{v}{\sqrt{2}}, \quad r \frac{d\theta}{dt} = \frac{v}{\sqrt{2}}$$

Obliczamy stosunek $\frac{dr}{d\theta}$:

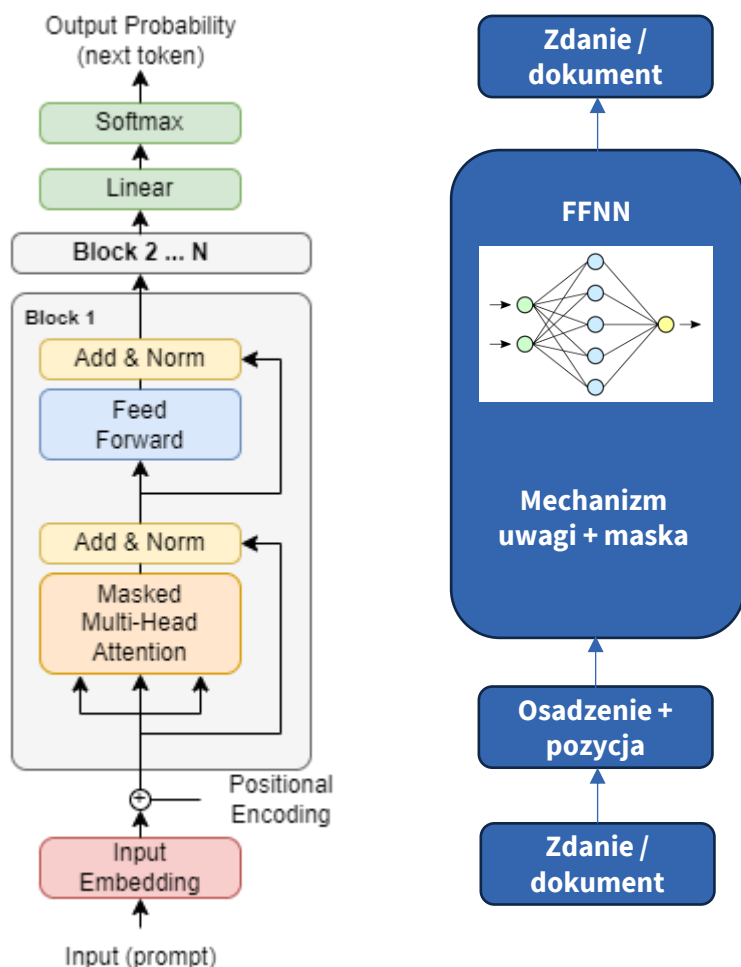
$$\frac{dr}{d\theta} = \frac{dr/dt}{d\theta/dt} = \frac{-\frac{v}{\sqrt{2}}}{\frac{v}{\sqrt{2}r}} = -r$$

Rozwiązujemy równanie różniczkowe:

$$\frac{dr}{r} = -d\theta$$

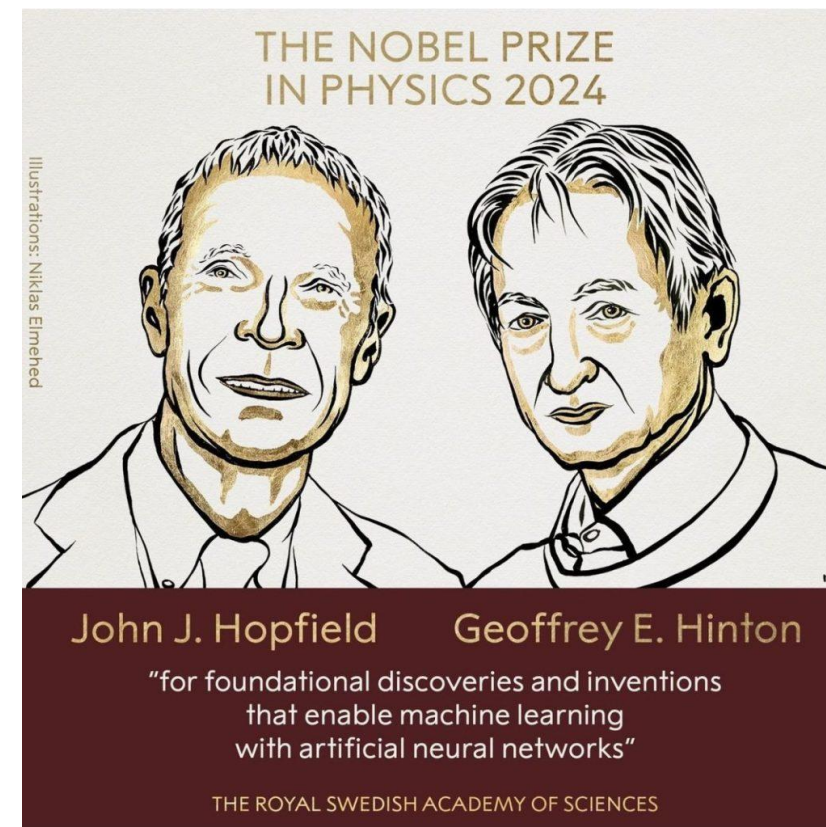
Całkujemy obustronnie:

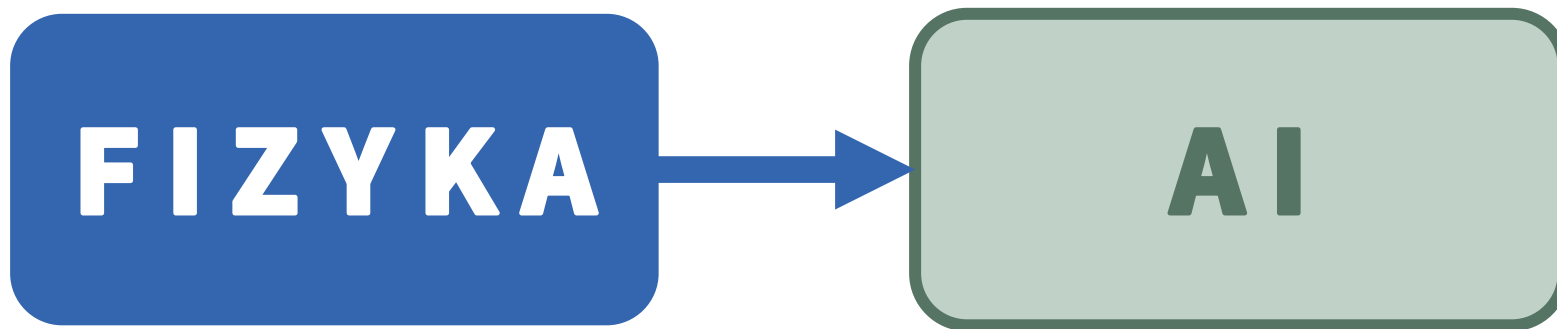
$$\int \frac{dr}{r} = - \int d\theta \Rightarrow \ln r = -\theta + C$$



... and the *cute* **kitten** *purred* and then ...
... the *cute* *furry* **cat** *purred* and *miaowed* ...
... that the *small* **kitten** *miaowed* and she ...
... the *loud* *furry* **dog** *ran* and *bit* ...

8 października 2024





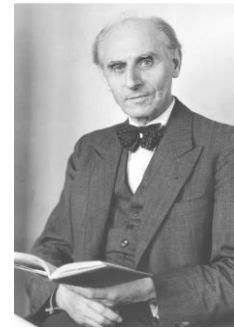


Model Isinga-Lenza (1920, 1924)

Prosty model ferromagnetyka

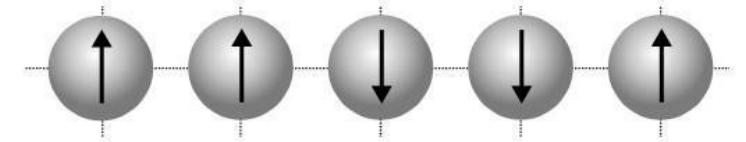


Ernst Ising



Wilhelm Lenz

Układ spinów $s_i = \pm 1$



$$s = \{+1, +1, -1, -1, +1\}$$

Przykładowa **konfiguracja spinów** s

Hamiltonian (energia) układu spinów s

$$H(s) = - \sum_{\langle ij \rangle} J_{ij} s_i s_j - \mu \sum_i h_i s_i$$

oddziaływanie
pomiedzy spinami (tylko
najbliżsi sąsiedzi)

zewnętrzne
pole

Prawdopodobieństwo konfiguracji s

$$P_{\beta}(s) = \frac{e^{-\beta H(s)}}{Z_{\beta}}$$

rozkład **Boltzmanna** aka rozkład **Gibbsa**
aka funkcja **softmax**

$$Z_{\beta} = \sum_s e^{-\beta H(s)}$$

Suma
statystyczna

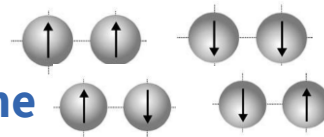
$$\beta = \frac{1}{k_B T}$$

Odwrotność
temperatury

$J_{ij} > 0$ – oddziaływanie **ferromagnetyczne**

$J_{ij} < 0$ – oddziaływanie **antyferromagnetyczne**

$J_{ij} = 0$ – **brak** oddziaływania



Brak pola

$$H(s) = - \sum_{\langle ij \rangle} J_{ij} s_i s_j$$

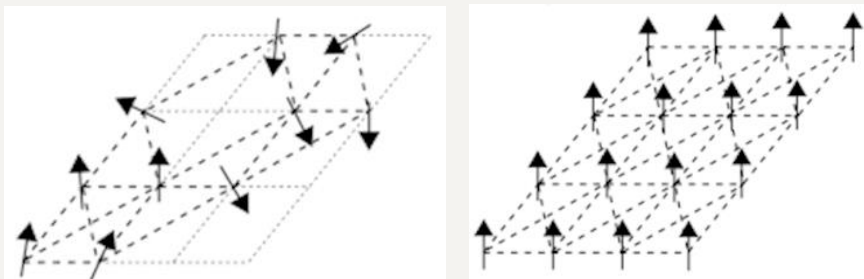
Jednorodne oddziaływania

$$H(s) = -J \sum_{\langle ij \rangle} s_i s_j$$



Szkło spinowe

Układ oddziałujących momentów magnetycznych – znak oddziaływania tych elementów jest losowy. Można to interpretować jako mieszankę połączeń ferromagnetycznych i antyferromagnetycznych.

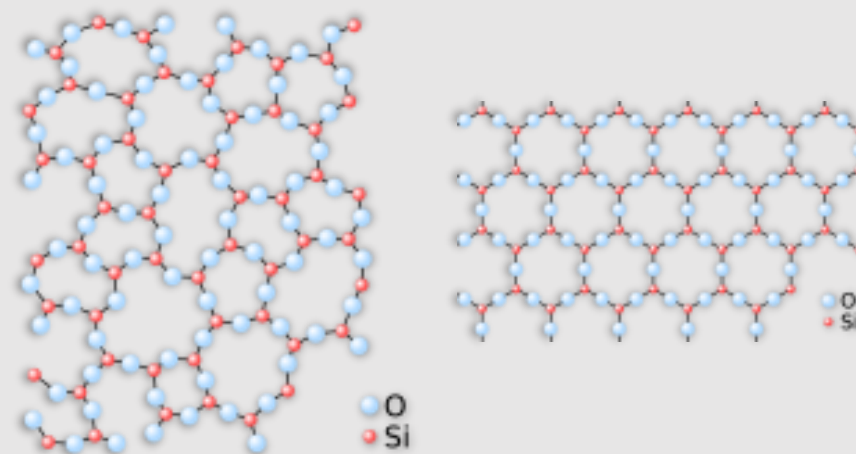


Szkło spinowe

Ferromagnetyk

Analogia

W zwykłym szkłe struktura wiązań jest nieregularna -> tak samo w szkłe spinowym w kontekście nieporządku magnetycznego w szkłe spinowym



Amorficzny SiO_2
(szkło kwarcowe)

Krystaliczny SiO_2
(kwarc)

Czy takie materiały w ogóle istnieją? Tak, np. stopy **złota z żelazem (AuFe)** czy **miedzi z manganem (CuMn)**



Model Edwardsa-Andersona

Bardzo podobny do modelu Isinga,

$$H(\mathbf{s}) = - \sum_{\langle ij \rangle} J_{ij} s_i s_j$$

ale wartości J_{ij} pochodzą z rozkładu Gaussa ze średnią J_0 i wariancją J^2

$$P(J_{ij}) = \sqrt{\frac{N}{2\pi J^2}} e^{-\frac{N}{2J^2} \left(J_{ij} - \frac{J_0}{N} \right)^2}$$

Model Sherringtona-Kirkpatricka

Hamiltonian prawie identyczny jak w przypadku modelu Edwardsa-Andersona

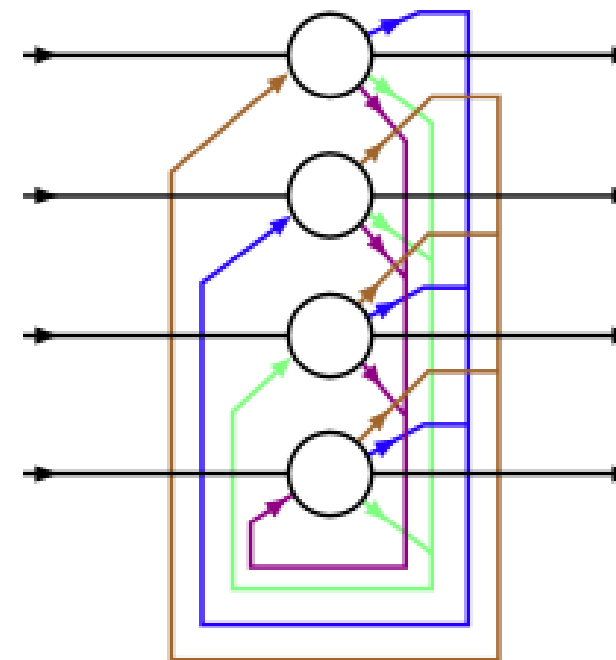
$$H(\mathbf{s}) = - \frac{1}{\sqrt{N}} \sum_{i < j} J_{ij} s_i s_j$$

Ale główna różnica wynika z wzięcia pod uwagę połączeń dalekozasięgowych.



Konstrukcja sieci Hopfielda (1982)

- Całkowicie połączona i symetryczna (wagi lub oddziaływania $w_{ij}=w_{ji}$) sieć
- Stany neuronów sieci Hopfielda przyjmują wartości binarne (neuron aktywny 1 / nieaktywny -1)
- Stan sieci opisujemy przez wektor $\mathbf{V}$ (N bitowe słowo) np. $V = \{+1, -1, +1, +1\}$ w przypadku sieci na rys. obok
- Oddziaływania pomiędzy neuronami w_{ij} przyjmują wartości +1 lub -1



Przykładowa sieć
4 neuronów

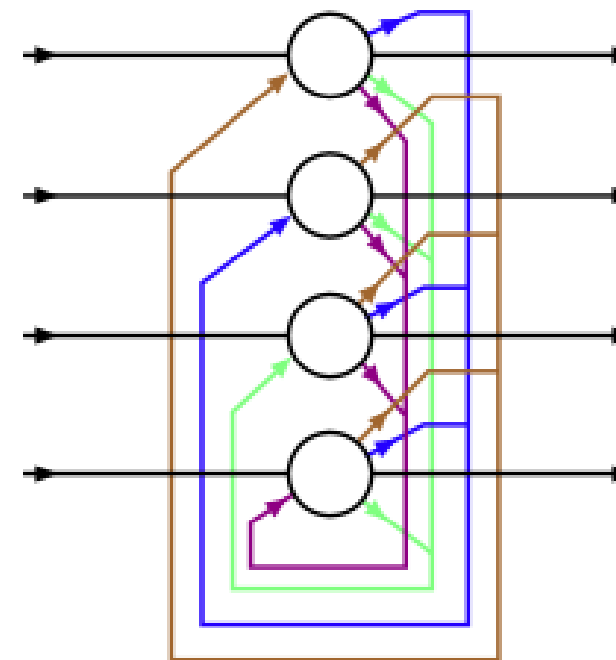


Uczenie się wzorca

- Załóżmy, że chcemy nauczyć sieć wzorca $V^s = \{+1, +1, -1, -1\}$
- Stosujemy tzw. **zasadę Hebba**, która mówi, że waga połączenia pomiędzy neuronami i oraz j to $w_{ij} = V_i^s V_j^s$
- Innymi słowy, na wejściu wstawimy V^s i czekamy aż sieć znajdzie stan stabilny (tzn ustalą się wartości w_{ij})
- Po tym jak sieć nauczy się wzorca, w_{ij} już się nie zmieniają

Odzyskiwanie wzorca

- Co się stanie, jeśli na wejściu wstawimy teraz jakiś inny wzorzec?
- Stan ten działa na sieć w ten sposób, że dokonujemy aktualizacji każdego neuronu jako 1 jeśli $\sum_j w_{ij} V_j > \theta_i$, -1 jeśli $\sum_j w_{ij} V_j < \theta_i$
- Stany neuronów będą się zmieniać do momentu aż “odzyskają” oryginalny wzorzec V^s



Przykładowa sieć
4 neuronów

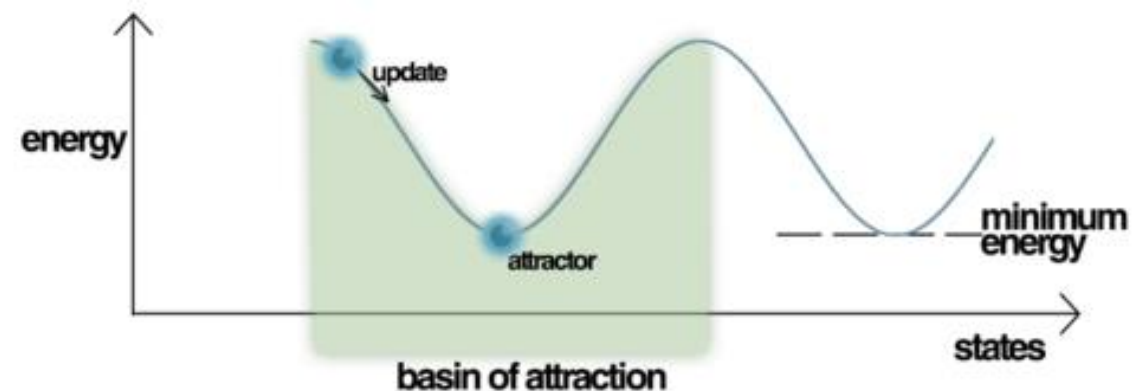


A co to ma wspólnego z modelem Isinga i szkłem spinowym?

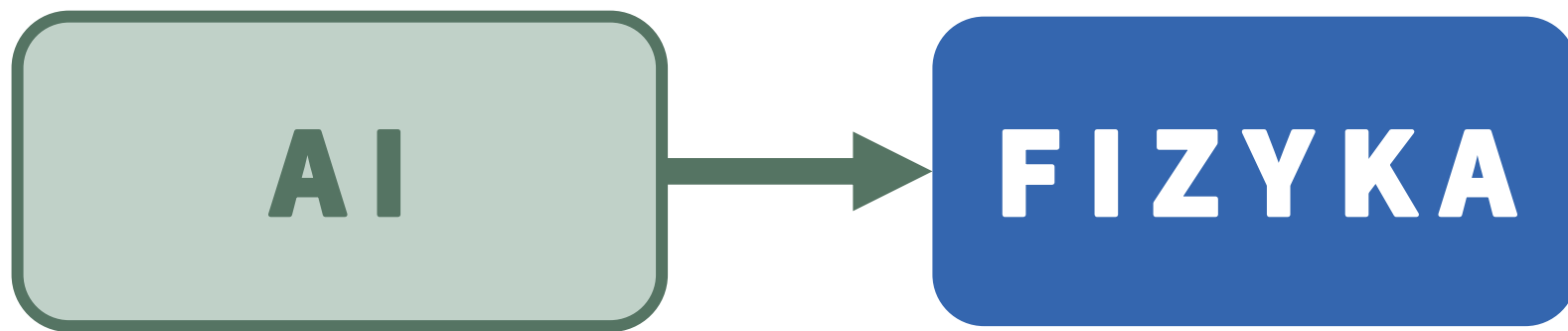
Z każdym stanem sieci Hopfielda można powiązać wartość skalarną, określaną jako “energię” tego stanu:

$$E = -\frac{1}{2} \sum_{i,j} w_{ij} s_i s_j - \sum_i \theta_i s_i$$

Wartość ta albo pozostaje na stałym poziomie albo też zmniejsza się podczas aktualizacji stanów neuronów.



Jak widać, opis sieci Hopfielda oraz szkła spinowego w modelu SK jest tożsamy (w przypadku losowych wag w SK). Zapamiętane wzorce odpowiadają lokalnym minimum energii.





Machine learning phases of strongly correlated fermions
Phys. Rev. X (2017)

Quantum loop topography for machine learning
Phys. Rev. Lett (2017)

Machine learning $\mathbb{Z}_2$ quantum spin liquids with quasiparticle statistics **Phys. Rev. B (2017)**

Learning phase transitions by confusion
Nat. Phys. (2017)

Machine learning topological invariants with neural networks
Phys. Rev. Lett (2018)

Extrapolating quantum observables with machine learning: inferring multiple phase transitions from properties of a single phase **Phys. Rev. Lett. (2018)**

Machine learning vortices at the Kosterlitz–Thouless transition **Phys. Rev. B 97 (2018)**

nature
physics

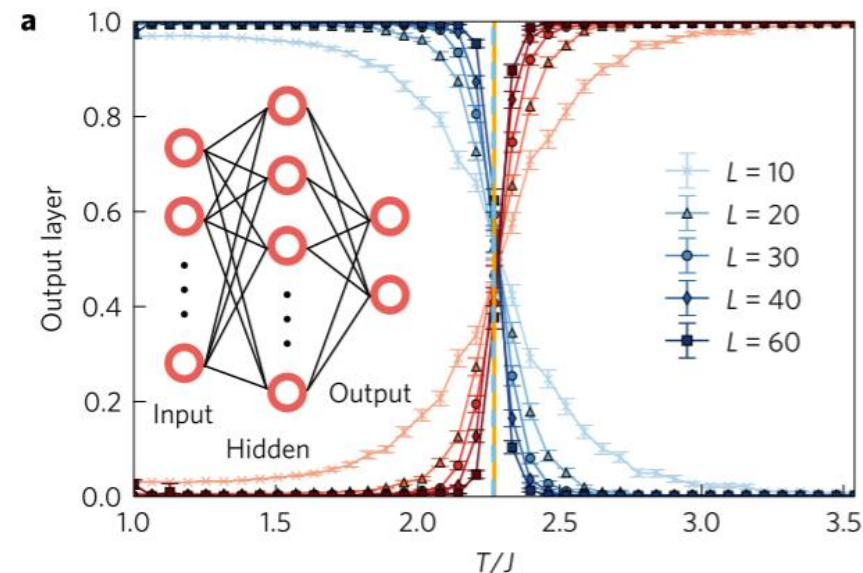
LETTERS

PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4035

Machine learning phases of matter

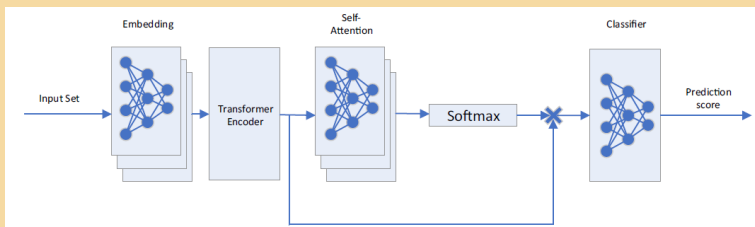
Juan Carrasquilla^{1*} and Roger G. Melko^{1,2}

- Uczenie pod nadzorem,
- Prosty 1D model Isinga

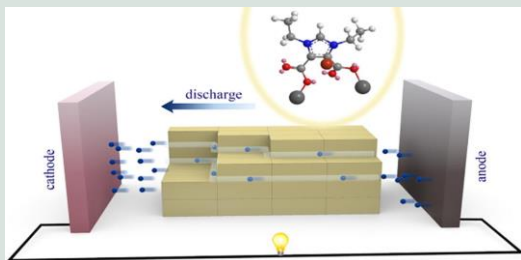




Identyfikacja cząstek przy braku danych

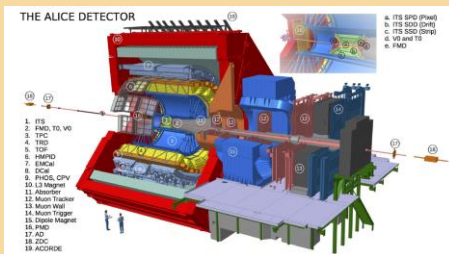


[M Kasak, K Deja, M Karwowska, M Jakubowska, Ł Graczykowski, M Janik, Eur. Phys. J. C (2024) 84:691]

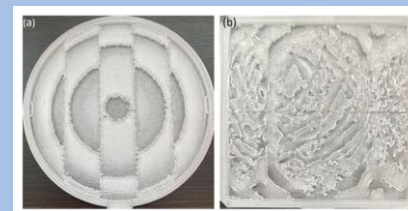
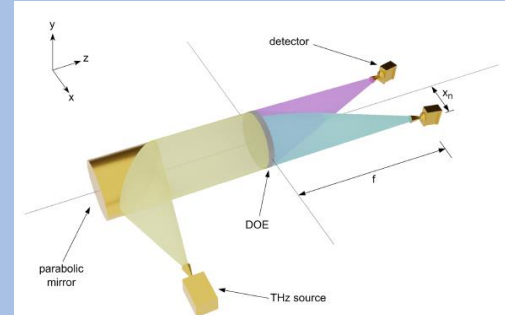


Wykorzystanie NN do badania przestrzeni międzyziarnowej w elektrolitach stałych

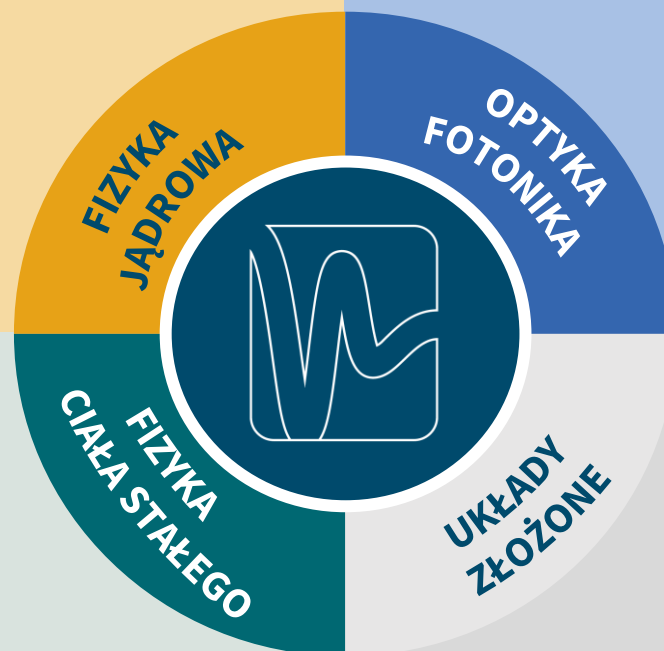
[M Kryński, Nowe podejście do modelowania komputerowego procesów zachodzących na interfejsach elektrolitów stałych, LAB-TECH of Excellence]



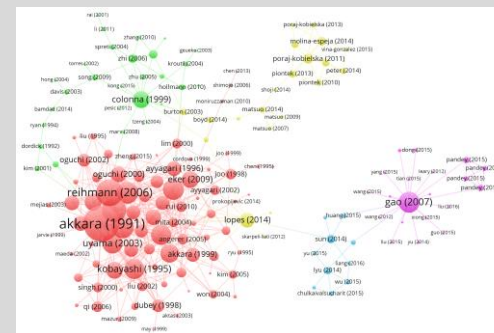
Wykorzystanie NN do projektowania dyfrakcyjnych elementów optycznych



[P Komorowski, M Kurowska, M Kaluza, P Czerwińska, P Zagrajek, A Siemion, IEEE Trans. Terahertz Sci. Technol. 14, 228 (2024)]



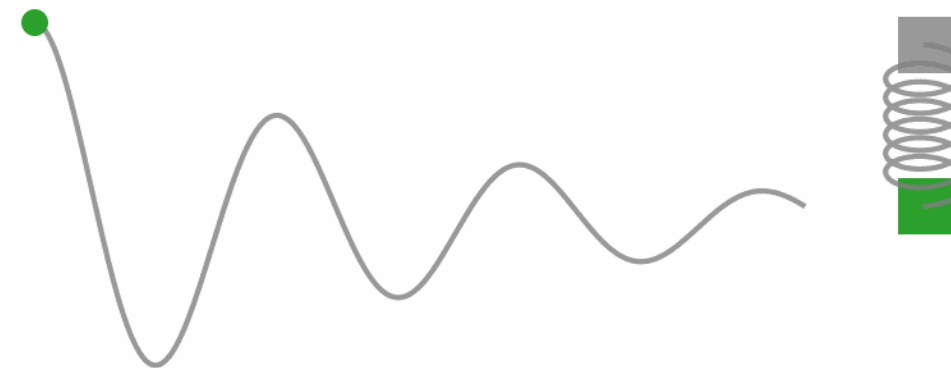
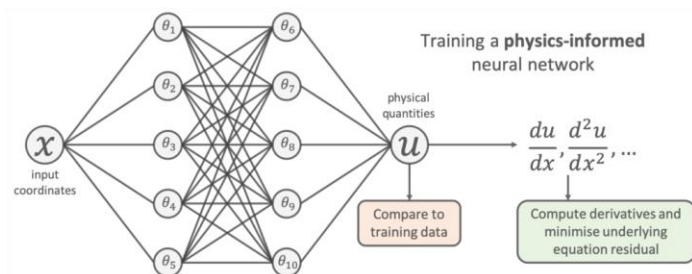
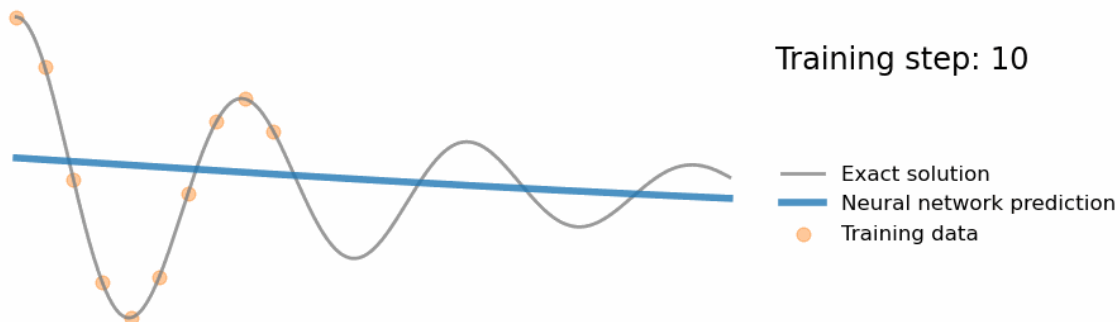
Wyjaśnialność przewidywań GNN w sieciach złożonych



[P Kaczyńska, J Sienkiewicz, D Ślęzak, Accumulated Local Effects for Link Prediction with Graph Neural Networks, submitted to NeurIPS 2024 Workshop ATTRIB]

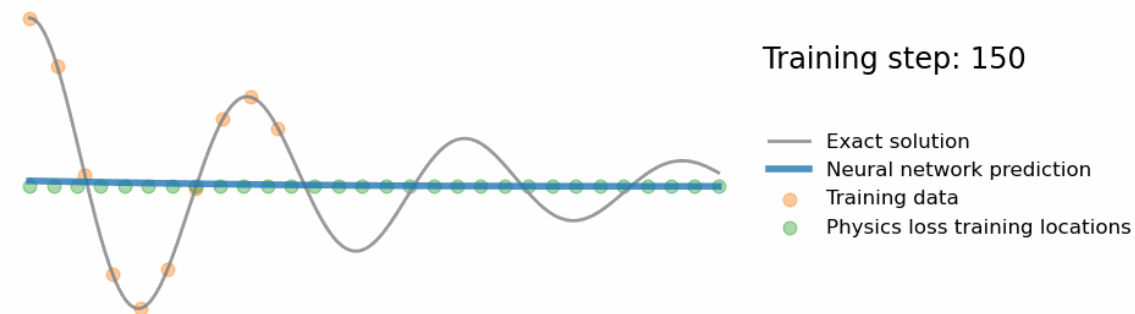


PINN Physics Informed Neural Networks



Tłumiony oscylator harmoniczny

Uwzględnienie więzów w funkcji kosztu





S. Lem, *Obłok Magellana*, 1955

Goobar to przysłuchiwał się odpowiedziom, to odczytywał je i powoli przeczył głową; czasem odwracał się ze zniechęceniem [...] zapoznawszy się ze wszystkim, co miała mu do zakomunikowania, powtarzał przeczący ruch głową i wypowiadał [...] nagle ściągnął brwi, [...] i znikł mi z oczu. [...] Po minucie Goobar wrócił; był z nim mechanoautomat, który skierował się prosto do mózgu elektrycznego, uczony cofnął się, zmrużył oko i przemówił do mechanoautomatu [...] wypuścił z głowicy narzędziowej wiertło i przewiercił dziurę w pancernym czole mózgu, po czym nożycową dźwignią odchylił całą pokrywę. [...] Goobar zaś z największym zaciekawieniem zaglądał do środka otwartej maszyny; potem wziął kilka drobnych instrumentów i począł zmieniać połączenia kabli; pracował niezwykle szybko. [...] Goobar usiadł na skraju wysokiego stołka i przez długi czas patrzył na rozwijające się w głębi ekranów krzywe, nareszcie potakująco kiwnął głową [...]. Pomyślałem, że przekształcił zapewne część aparatury aksjomatycznej; widać tworzył nie istniejącą gałąź matematyki, której konieczność pojawiła się w związku z nowymi dociekaniem, i byłem świadkiem operacji, jaką wprowadzał rozumowanie mózgu na nowe tory.





SCIENCE ADVANCES | RESEARCH ARTICLE

COMPUTER SCIENCE

AI Feynman: A physics-inspired method for symbolic regression

Silviu-Marian Udrescu¹ and Max Tegmark^{1,2*}

A core challenge for both physics and artificial intelligence (AI) is symbolic regression: finding a symbolic expression that matches data from an unknown function. Although this problem is likely to be NP-hard in principle, functions of practical interest often exhibit symmetries, separability, compositionality, and other structures. In this spirit, we develop a recursive multidimensional symbolic regression algorithm that discovers all of them, while previous publicly available software cracks only a subset. On a standard test set, we improve the state-of-the-art success rate from 15 to 90%.

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
License.

nature communications

Article

<https://doi.org/10.1038/s41467-024-50074-w>

Evolving scientific discovery by unifying data and background knowledge with AI Hilbert

Received: 23 September 2023

Ryan Cory-Wright¹, Cristina Cornelio², Sanjeeb Dash³, Bachir El Khadir³ & Lior Hershkov³

PNAS

PERSPECTIVE

Learning dynamical systems from data: An introduction to physics-guided deep learning

Rose Yu^{a,1} and Rui Wang^{a,b}

Edited by Yuhai Tu, International Business Machines Corp, Yorktown Heights, NY; received August 28, 2023; accepted November 21, 2023, by Editorial Board Member Herbert Levine

OPEN ACCESS



simultaneously explain natural phenomena. A key goal in science is to discover laws by manipulating equations, and verifying them against experimental data within the model. We propose a solution to this problem: we express the model as polynomials and model notions of minimal constraints, solve polynomial optimization, and solve in a principled manner using the fact that some famous scientific theories, such as the Radiated Gravitational waves, are explained in a principled manner from

nature computational science

Article

<https://doi.org/10.1038/s41467-024-50074-w>

Constructing custom thermodynamics using deep learning

Received: 13 August 2023

Accepted: 30 November 2023

Published online: 29 December 2023

Check for updates

Xiaoli Chen^{1,2*}, Beatrice W. Soh^{3,4}, Zi-En Ooi⁵, Eleonore Vissol-Gaudin⁶, Haljun Yu^{1,2*}, Kostya S. Novoselov^{2,3,4}, Kedar Hippalgaonkar^{2,3,4} & Qianxiao Li^{1,2*}

One of the most exciting applications of artificial intelligence is automated scientific discovery based on previously amassed data, coupled with restrictions provided by known physical principles, including symmetries and conservation laws. Such automated hypothesis creation and verification can assist scientists in studying complex phenomena, where traditional physical intuition may fail. Here we develop a platform based on a generalized Onsager principle to learn macroscopic dynamical descriptions of arbitrary stochastic dissipative systems directly from observations of their microscopic trajectories. Our method simultaneously constructs reduced thermodynamic coordinates and interprets the dynamics on these coordinates. We demonstrate its effectiveness by studying theoretically

Chaos, Solitons and Fractals 188 (2024) 115538



Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

journal homepage: www.elsevier.com/locate/chaos



AI-Lorenz: A physics-data-driven framework for Black-Box and Gray-Box identification of chaotic systems with symbolic regression

Mario De Florio^{a,*}, Ioannis G. Kevrekidis^{c,d}, George Em Karniadakis^{a,b}

^a Division of Applied Mathematics, Brown University, Providence, RI, USA

^b School of Engineering, Brown University, Providence, RI, USA

^c Department of Chemical and Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, USA

^d Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, USA



SCIENCE ADVANCES | RESEARCH ARTICLE

COMPUTER SCIENCE

AI Feynman: A physics-inspired method for symbolic regression

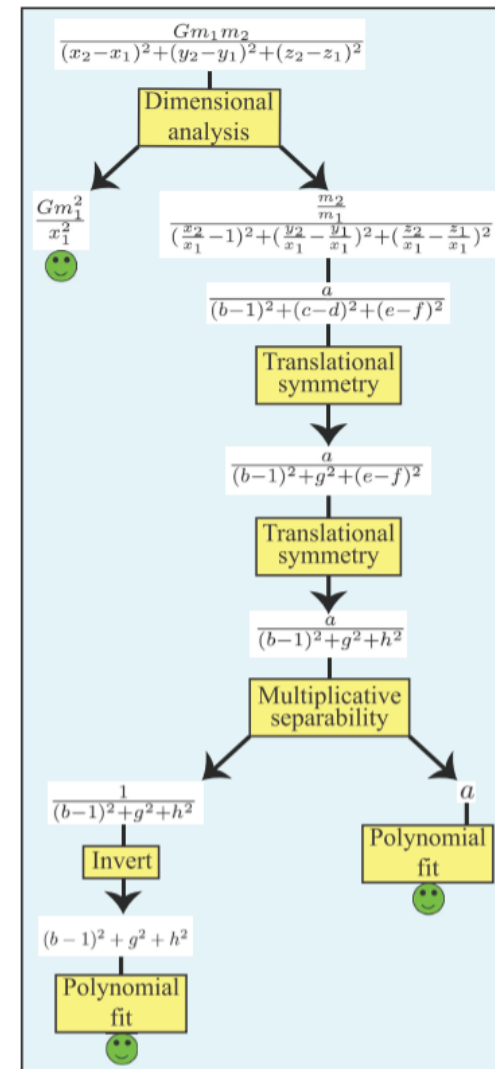
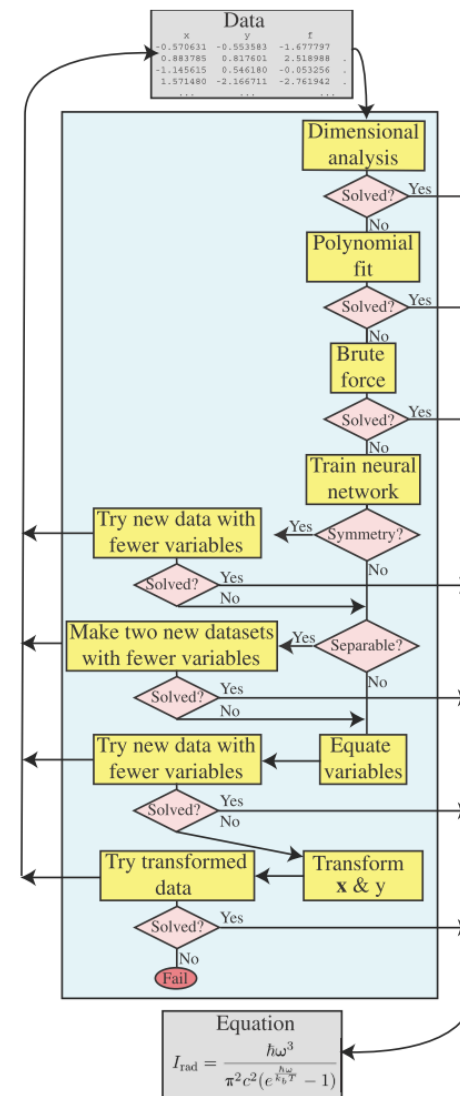
Silviu-Marian Udrescu¹ and Max Tegmark^{1,2*}

A core challenge for both physics and artificial intelligence (AI) is symbolic regression: finding a symbolic expression that matches data from an unknown function. Although this problem is likely to be NP-hard in principle, functions of practical interest often exhibit symmetries, separability, compositionality, and other simplifying properties. In this spirit, we develop a recursive multidimensional symbolic regression algorithm that combines neural network fitting with a suite of physics-inspired techniques. We apply it to 100 equations from the *Feynman Lectures on Physics*, and it discovers all of them, while previous publicly available software cracks only 71; for a more difficult physics-d test set, we improve the state-of-the-art success rate from 15 to 90%.

Copyright © 2020
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

Ogólnie mamy dane w postaci $\{x_1, x_2, \dots, x_n, y\}$ gdzie $y=f(x_1, x_2, \dots, x_n)$ I szukamy f:

- (1) Stosując rachunek jednostek
- (2) Szukając wielomianów niskiego rzędu,
- (3) Zakładając, że f składa się z funkcji elementarnych,
- (4) f jest ciągła,
- (5) f cechuje się symetrią,
- (6) f jest separowalna





nature computational science

Article

<https://doi.org/10.1038/s43588-023-00581-5>

Constructing custom thermodynamics using deep learning

Received: 13 August 2023

Accepted: 30 November 2023

Published online: 29 December 2023

Check for updates

Xiaoli Chen^{1,2,7}, Beatrice W. Soh^{3,7}, Zi-En Ooi³, Eleonore Vissol-Gaudin¹,
Haljun Yu^{1,6}, Kostya S. Novoselov^{2,3,4}, Kedar Hippalgaonkar^{1,2,3,4} &
Qianxiao Li^{1,2,3,4}

One of the most exciting applications of artificial intelligence is automated scientific discovery based on previously amassed data, coupled with restrictions provided by known physical principles, including symmetries and conservation laws. Such automated hypothesis creation and verification can assist scientists in studying complex phenomena, where traditional physical intuition may fail. Here we develop a platform based on a generalized Onsager principle to learn macroscopic dynamical descriptions of arbitrary stochastic dissipative systems directly from observations of their microscopic trajectories. Our method simultaneously constructs reduced thermodynamic coordinates and interprets the dynamics on these coordinates. We demonstrate its effectiveness by studying theoretically

Opierając się na relacjach Onsagera, oprogramowanie S-Onsager-Net tworzy zakmnięty zestaw współrzędnych, opisujących układ.

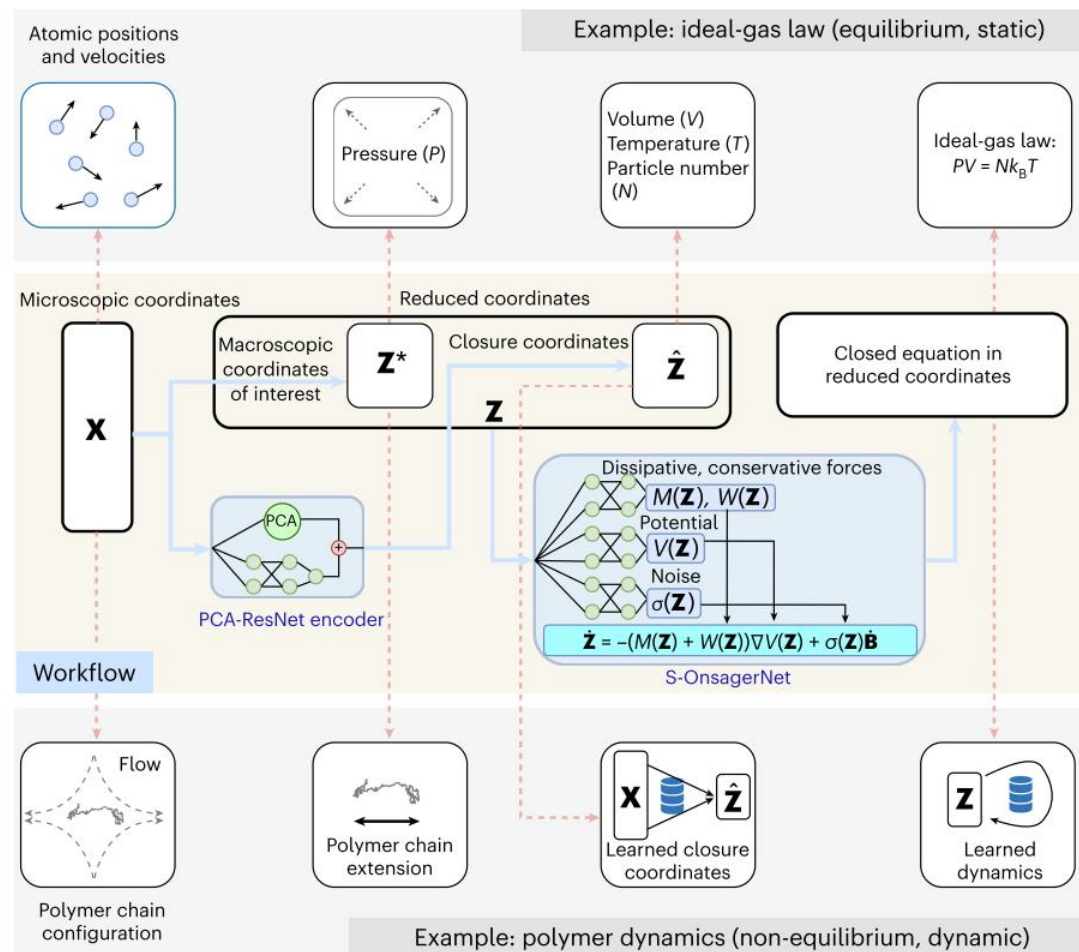
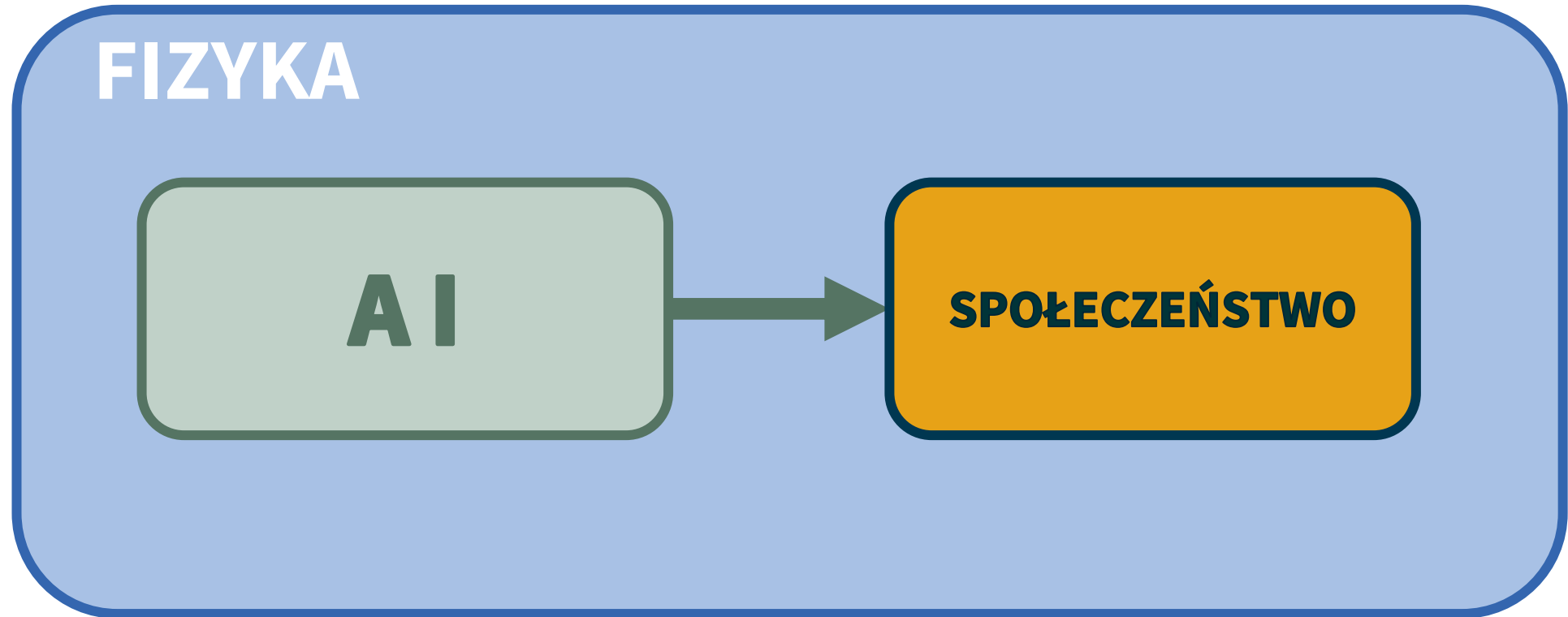


Fig. 1 | Overall workflow of the proposed approach. Given a complex system described by $\mathbf{X}$, the goal is to model the behavior of macroscopic coordinates of interest $\mathbf{Z}^*$. We construct closure coordinates $\hat{\mathbf{Z}}$ and closed (dynamical) equation on the combined reduced coordinates $\mathbf{Z} = (\mathbf{Z}^*, \hat{\mathbf{Z}})$. The classical ideal-gas law is an illustration of this process (top). For general non-equilibrium, dynamic systems (bottom), carrying out this workflow from theoretical analysis is

challenging. Our machine learning method (middle) addresses this by simultaneously constructing the closure coordinates using PCA-ResNet (Methods), and governing equations on reduced coordinates using the S-OnsagerNet with drift term $-(M(\mathbf{Z}) + W(\mathbf{Z})) \nabla V(\mathbf{Z})$ and noise term $\sigma(\mathbf{Z})$ (see equation (2)).





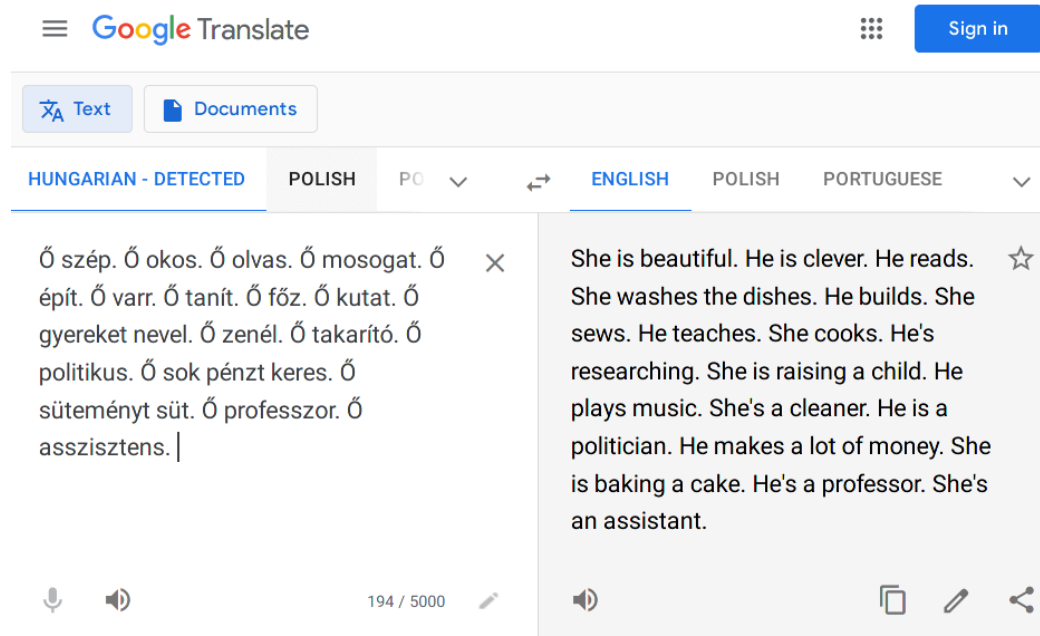
**Wzmacnianie
stereotypów**

**Przeciążenie
informacyjne**

**Tworzenie baniek
informacyjnych**

**Homogenizacja
przestrzeni informacyjnej**

**Polaryzacja
społeczeństwa**



TAY – Thinking About You

“The AI with zero chill” –
chatbot (2016)



**Wzmacnianie
stereotypów**

**Przeciążenie
informacyjne**

**Tworzenie baniek
informacyjnych**

**Homogenizacja
przestrzeni informacyjnej**

**Polaryzacja
społeczeństwa**

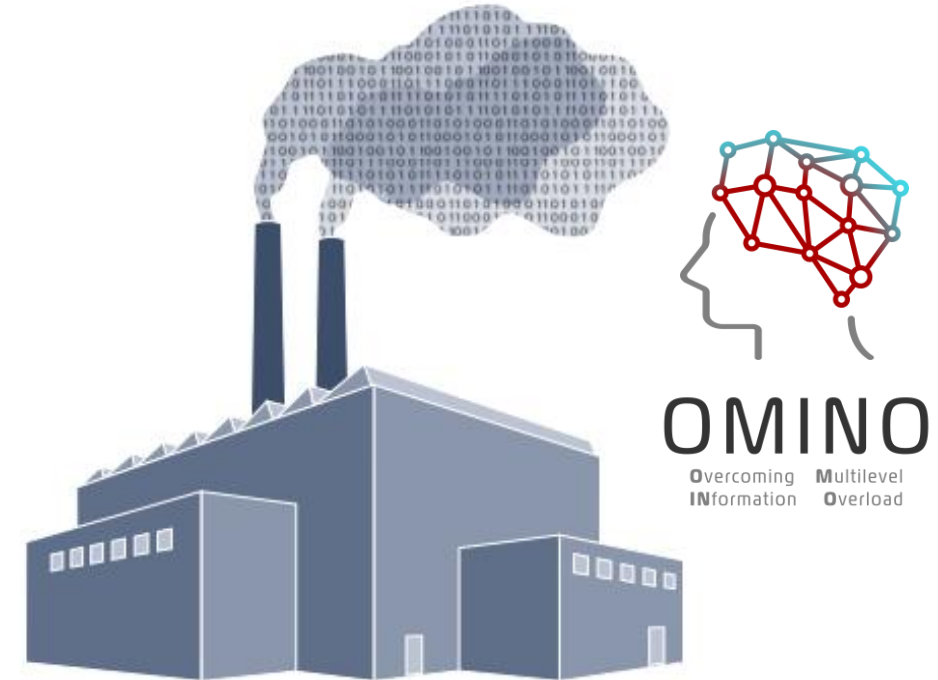


Fig. 1 | Information space should be considered as an important part of our environment and be protected alongside air, water and soil. Smoking

Koszt produkcji informacji ≈ 0



Wzmacnianie
stereotypów

Przeciążenie
informacyjne

Tworzenie baniek
informacyjnych

Homogenizacja
przestrzeni informacyjnej

Polaryzacja
społeczeństwa

nature machine intelligence

Article

<https://doi.org/10.1038/s42256-023-00731-4>

Human–AI adaptive dynamics drives the emergence of information cocoons

Received: 30 March 2023

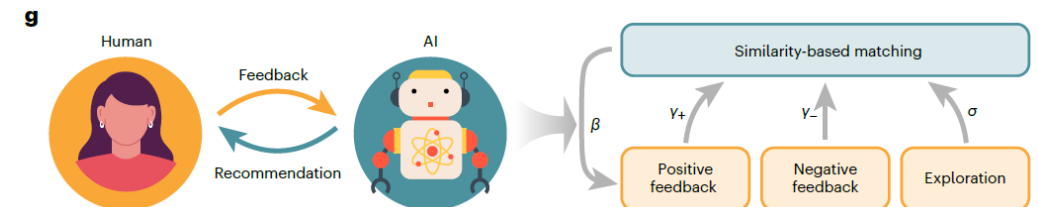
Accepted: 6 September 2023

Published online: 9 October 2023

Check for updates

Jinghua Piao^{1,2}, Jiazhen Liu^{1,2}, Fang Zhang², Jun Su² & Yong Li¹✉

Despite AI-driven recommendation algorithms being widely adopted to counter **information overload**, substantial evidence suggests that they are building cocoons of homogeneous contents and viewpoints, further aggravating **social polarization and prejudice**. Curbing these perils requires a deep insight into the origin of information cocoons. Here we investigate information cocoons in the real world using two large datasets and find that a large number of users are trapped in **information cocoons**. Further empirical analysis suggests that two ingredients, each corresponding to a fundamental mechanism in human–AI interaction systems, are correlated with the **loss of information diversity**. Grounded on the empirical findings,



$$d\mathbf{u}_t = \gamma_+ \sum_{\mathbf{i}_k \in R^+_\beta} F(\mathbf{u}_t, \mathbf{i}_k) dt + \gamma_- \sum_{\mathbf{i}_k \in R^-_\beta} F(\mathbf{u}_t, \mathbf{i}_k) dt + \sigma dW_t$$



Wzmacnianie
stereotypów

Przeciążenie
informacyjne

Tworzenie baniek
informacyjnych

Homogenizacja
przestrzeni informacyjnej

Polaryzacja
społeczeństwa

PNAS

RESEARCH ARTICLE | SOCIAL SCIENCES

Link recommendation algorithms and dynamics of polarization in online social networks

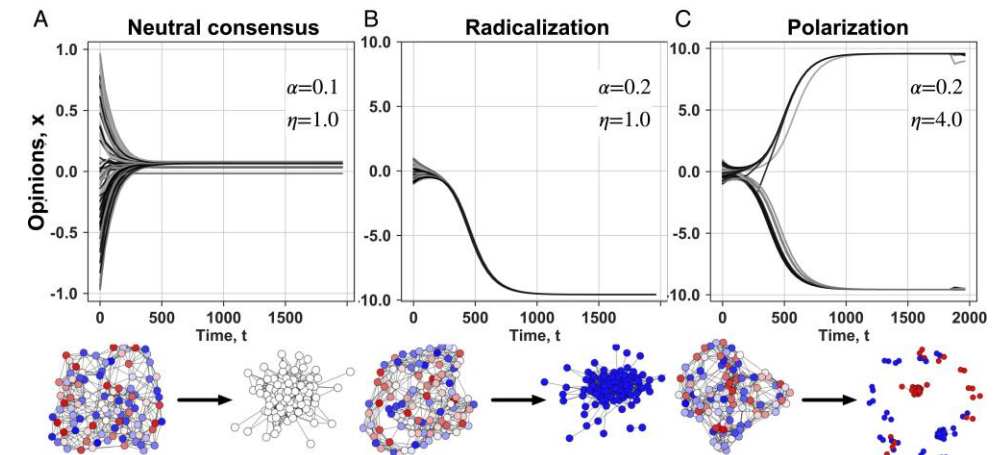
Fernando P. Santos ^{a,b,1}, Yphtach Lelkes ^c, and Simon A. Levin ^a

Edited by Charles Perrings, Arizona State University, Tempe, AZ, and accepted by Editorial Board Member Kenneth W. Wachter July 14, 2021 (received for review March 1, 2021)

December 6, 2021 | 118 (50) e2102141118 | <https://doi.org/10.1073/pnas.2102141118>

$$x_i(t+1) = \gamma x_i(t) + K \sum_j^N A_{i,j} \tanh(\alpha x_j(t)) / k_i$$

$$s_{i,j} = |N_i \cap N_j|^\eta / \sum_k^N |N_i \cap N_k|^\eta$$





Comment

<https://doi.org/10.1038/s42254-024-00772-7>

Physics and the empirical gap of trustworthy AI

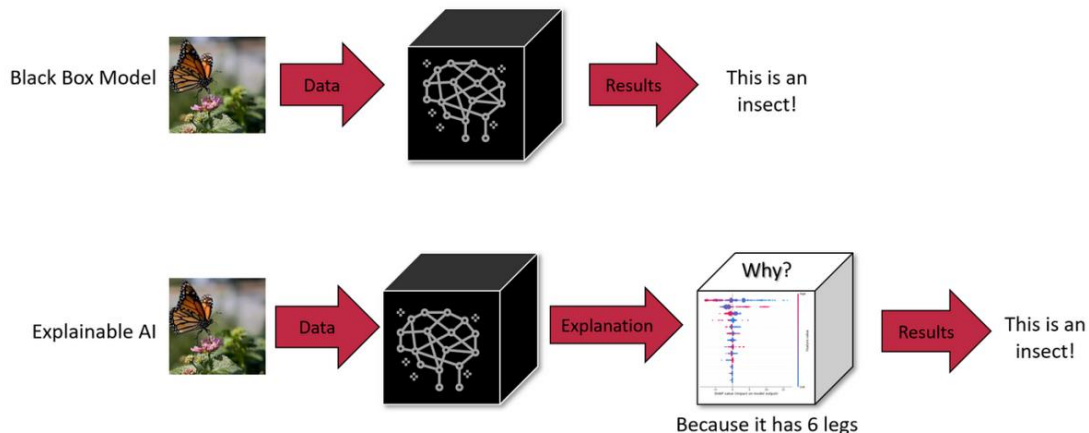
Savannah Thais

Check for updates

Understanding what cutting-edge AI models are doing ‘under the hood’ requires not just theoretical research but also well-controlled computational experiments. Savannah Thais explains why physics datasets may be the testing ground that AI developers need and how physicists can play a critical role in developing trustworthy AI.

reliable knowledge about the system it is observing, it is necessary for the measurement and evaluation process to be well characterized and understood; for example, in astronomy research, it is critical that we understand precisely the performance and uncertainty of our telescopes.

I believe that physics can be uniquely useful to this process of measuring AI model performance and reliability. With data drawn from physics experiments or observations of physical systems, we have a (reasonably) robust mathematical understanding of the true data generating functions, as well as the symmetries and conservation laws obeyed by the data; this stands in contrast to many other types



Rozgryzamy AI.

Centrum Wiarygodnej Sztucznej Inteligencji
Politechnika Warszawska

https://credibleai.github.io/job_all.html

Open Positions in Centre for Credible AI



Dziękuję
za uwagę