

**Referat wygłoszony w dniu 22.05.2025 roku
podczas posiedzenia naukowego Komisji Ergonomicznej PAN**

dr hab. inż. Maciej Jaworski, prof. PK
Politechnika Krakowska
Katedra Informatyki
mail: maciej.jaworski@pk.edu.pl
ORCID: orcid.org/my-orcid?orcid=0000-0002-8410-4231



AI Wczoraj i Dziś

Sztuczna inteligencja (AI) to dziedzina informatyki zajmująca się tworzeniem systemów zdolnych do samodzielnego uczenia się, adaptacji i rozwiązywania problemów w sposób przypominający ludzkie myślenie. Różne definicje inteligencji są niezbyt ścisłe i podkreślają różne jej aspekty: zdolność do działania celowego, rozumowania czy skutecznego radzenia sobie z otoczeniem. Inteligencja bywa także opisywana jako wynik procesu przetwarzania, porównywania i wykorzystywania informacji w nowych kontekstach. W niektórych koncepcjach pojawia się też pogląd, że inteligencję można rozumieć jako efektywność w kompresji informacji, co ma bezpośrednie zastosowanie w rozwoju algorytmów uczących się. Przykładem takiego podejścia jest zadanie kompresji dużych zbiorów danych tekstowych przy minimalnym rozmiarze kodu programu i archiwum wynikowego, co może odzwierciedlać zdolność algorytmu do uogólniania i "rozumienia" treści.

Istotnym podzbiorem metod sztucznej inteligencji stanowi uczenie maszynowe, w którym kluczowym elementem jest zdolność systemu do samodzielnego pozyskiwania wiedzy z danych, bez konieczności ręcznego programowania zasad postępowania. Wyróżnia się trzy podstawowe typy uczenia maszynowego. W uczeniu nadzorowanym system trenowany jest na danych wejściowych wraz z przypisanymi etykietami, co pozwala mu nauczyć się przewidywać wyniki dla nowych, nieznanych danych. Typowe zastosowania obejmują klasyfikację i regresję. Uczenie nienadzorowane działa bez etykiet i koncentruje się na znajdowaniu ukrytej struktury w danych, jak grupowanie czy redukcja wymiarowości. Natomiast uczenie przez wzmacnianie opiera się na podejmowaniu decyzji na podstawie informacji o stanie środowiska oraz systemie nagród i kar, co pozwala wykształcić strategię maksymalizującą korzyści w dłuższym horyzoncie czasowym. Przykładem zadania, do którego dobrze pasuje paradygmat uczenia przez wzmacnianie, jest uczenie algorytmu grania w różne gry, np. szachy. Różne metody uczenia maszynowego – od prostych modeli liniowych po złożone algorytmy

zespołowe – oferują szeroki wachlarz narzędzi analizy danych i predykcji. Są to m.in. różne metody regresji, metody bayesowskie, drzewa decyzyjne czy też sieci neuronowe. Te ostatnie stanowią obecnie najbardziej popularne narzędzie uczenia maszynowego, dające najbardziej efektywne wyniki dla wielu trudnych zadań, takich jak rozumienie obrazów czy przetwarzanie języka naturalnego.

Sieci neuronowe [1] to struktury inspirowane funkcjonowaniem biologicznego mózgu. Każdy sztuczny neuron otrzymuje sygnały z wielu wejść, sumuje je z odpowiednimi wagami i przekształca wynik za pomocą funkcji aktywacji. Uczenie polega na modyfikowaniu wartości tych wag, tak aby uzyskać pożądaną odpowiedź na dane wejściowe. Pierwsze sieci, jak perceptron Rosenblatta (1958) [2], miały bardzo ograniczone możliwości, jednak rozwój algorytmu wstecznej propagacji błędów oraz postęp technologiczny umożliwiły tworzenie coraz głębszych i bardziej wydajnych modeli. Dzięki metodom gradientowym sieci mogą uczyć się nawet bardzo złożonych zależności, pod warunkiem odpowiedniego doboru architektury, funkcji straty i strategii optymalizacji. Obecnie sieci neuronowe są powszechnie wykorzystywane w wielu dziedzinach, gdzie konieczne jest przetwarzanie dużych ilości danych o nieliniowych zależnościach.

W obszarze przetwarzania obrazów szczególne znaczenie mają konwolucyjne sieci neuronowe (CNN), które dzięki lokalnym filtrom i warstwom agregującym są wyjątkowo skuteczne w analizie danych wizualnych. Systemy tego typu osiągają wysoką skuteczność w zadaniach klasyfikacji obrazów, detekcji obiektów oraz segmentacji semantycznej. Rozwój benchmarkowych zbiorów danych, takich jak ImageNet, pozwolił na porównywanie i doskonalenie architektur, czego przykładem są sieci AlexNet, VGG, ResNet czy GoogLeNet. Złożoność współczesnych modeli sięga miliardów parametrów, a ich trenowanie wymaga ogromnych zbiorów danych oraz specjalizowanych jednostek obliczeniowych. Zastosowania obejmują m.in. diagnostykę medyczną, gdzie algorytmy potrafią przewyższać lekarzy w wykrywaniu niektórych patologii, znacząco redukując liczbę fałszywych wyników i poprawiając wskaźniki diagnostyczne.

Głębokie sieci neuronowe stanowią także podstawę współczesnych metod przetwarzania języka naturalnego, jakimi są wielkie modele językowe, takie jak GPT. Oparte są one na architekturze transformera, której kluczowym komponentem jest mechanizm uwagi (attention) [3]. Pozwala on analizować kontekst w długich sekwencjach tekstu i skutecznie modelować zależności między słowami. Modele te są trenowane na dziesiątkach terabajtów danych tekstowych i zawierają setki miliardów parametrów [4]. Dzięki temu są w stanie generować spójne i logiczne wypowiedzi, tłumaczyć teksty, streszczać artykuły czy prowadzić konwersacje w języku naturalnym. Mimo że często określane są mianem "papug stochastycznych", ze względu na statystyczny charakter predykcji, ich zdolność do przechodzenia testu Turinga i symulowania zrozumienia stawia je w centrum debat nad przyszłością AI. Koszty trenowania takich modeli sięgają dziesiątek milionów dolarów, co sprawia, że dostęp do tej technologii jest ograniczony do najbogatszych firm i instytucji.

Generatywna sztuczna inteligencja to obszar, w którym systemy uczą się tworzyć nowe dane – teksty, obrazy, muzykę, a nawet filmy – bazując na przykładach ze zbiorów uczących. Modele GAN (Generative Adversarial Networks) [5] składają się z dwóch komponentów: generatora

tworzącego dane i dyskryminatora oceniającego ich jakość, co prowadzi do iteracyjnego doskonalenia wyników. Modele dyfuzyjne [6] generują obrazy, zaczynając od losowego szumu i stopniowo go "odslaniając" w procesie przypominającym filtrowanie zakłóconego sygnału. Dzięki pracy w przestrzeni ukrytej, zamiast bezpośrednio na pikselach, osiągają bardzo wysoką jakość generowanych obrazów. Narzędzia takie jak MidJourney potrafią przekształcić krótkie opisy tekstowe (tzw. „prompty”) w szczegółowe ilustracje, a bardziej zaawansowane systemy – jak SORA – generują dość spójne wideo, starając się uwzględnić ciągłość ruchu, fizykę i perspektywę. Kluczowe wyzwania obejmują utrzymanie detali oraz spójność między klatkami, co wymaga znacznych zasobów obliczeniowych i dopracowanych algorytmów.

Zastosowanie AI nie ogranicza się jedynie do obrazów czy tekstów. W biologii molekularnej przełomem było stworzenie narzędzia AlphaFold – systemu potrafiącego przewidywać trójwymiarową strukturę białek na podstawie ich sekwencji aminokwasowej. Dzięki niemu możliwe stało się stworzenie kompletnej bazy struktur białkowych człowieka oraz milionów innych organizmów. Dotychczasowe metody były czasochłonne i kosztowne, natomiast AI pozwoliło na automatyzację i znaczne przyspieszenie tego procesu. Skala osiągnięcia została doceniona przyznaniem nagrody Nobla z chemii w roku 2024, a sam projekt wyznaczył nowy standard w zastosowaniach sztucznej inteligencji w naukach przyrodniczych. Nie była to jedyna nagroda Nobla przyznana w roku 2024 za osiągnięcia w dziedzinie sztucznej inteligencji – John Hopfield i Geof Hinton otrzymali nagrodę Nobla z fizyki odpowiednio za opracowanie sieci Hopfielda i maszyn Boltzmanna.

Przyszłość sztucznej inteligencji pozostaje niepewna, jednak skala inwestycji oraz tempo rozwoju technologii sugerują, że kolejne dekady przyniosą dalsze przełomy. Coraz częściej pojawia się koncepcja Artificial General Intelligence (AGI) – systemów, które nie tylko specjalizują się w jednym zadaniu, ale potrafią uczyć się, rozumować, analizować wielomodalne dane i działać elastycznie w różnych dziedzinach. Tego rodzaju AI mogłaby stać się równorzędnym partnerem człowieka w pracy intelektualnej, ale rodzi to też istotne pytania etyczne, społeczne i regulacyjne. Rozwój technologii powinien iść w parze z refleksją nad jej konsekwencjami, tak aby inteligencja – sztuczna czy nie – służyła dobru wspólnemu.

Bibliografia

- [1] Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep Learning*. MIT Press.
- [2] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.
- [3] Vaswani, A., et al. (2017). Attention is all you need. *31st Conference on Neural Information Processing Systems (NIPS)*. *Advances in Neural Information Processing Systems*, 30.
- [4] Brown T. B., et al. (2020). Language models are few-shot learners. *arXiv:2005.14165*.
- [5] Goodfellow, I., et al. (2014). Generative Adversarial Nets. *28th Conference on Neural Information Processing Systems (NIPS)*. *Advances in Neural Information Processing Systems*, pp. 2672–2680.
- [6] Prince, S. J. D. (2024). *Understanding Deep Learning*. MIT Press.