

Historia sztucznej inteligencji w medycynie

prof. Bartłomiej W. Papież, PhD, MSc

Big Data Institute, University of Oxford, Wielka Brytania

Skróty: AI (*artificial intelligence*) – sztuczna inteligencja; CNN (*convolutional neural networks*) – splotowe (konwolucyjne) sieci neuronowe; GPU (*graphics processing unit*) – karta graficzna

Wszystkie ilustracje zostały wygenerowane za pomocą Microsoft Copilot.

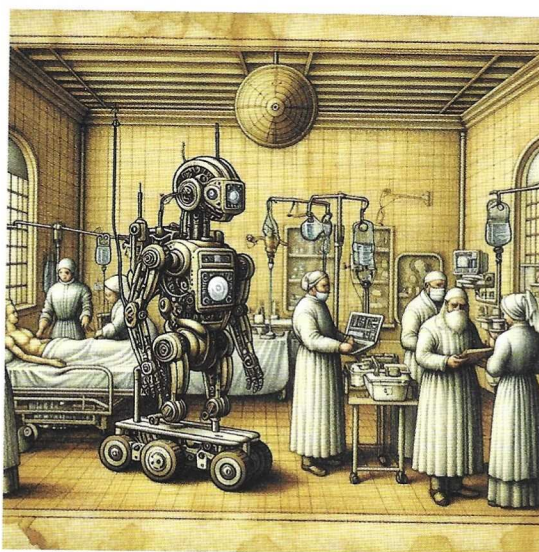
Obejrzyj wykład prof. B. Papieża „Sztuczna inteligencja w medycynie – aktualne zastosowania i perspektywy”



„Młody człowieku” – słyszę gdzieś za plecami, ale mój mózg już od jakiegoś czasu nie zwraca uwagi na takie zawołanie, a ja nadal konwersuję z jednym z uczestników konferencji o tym, co może, a czego jeszcze nie może sztuczna inteligencja (AI) za nas zrobić. „Młody człowieku, mam tylko jedno pytanie” – dochodzi do mnie po raz kolejny i tym razem już czuję, że ktoś mnie woła. „Czy sztuczna inteligencja zastąpi mnie w przychodni? Bo ja mam jeszcze prawie 5 lat do emerytury!”

Czy maszyna może myśleć?

Trudno określić, kiedy się narodziła idea sztucznej inteligencji. Prób opisów i stworzenia myślących maszyn pomagających człowiekowi można się doszukiwać w czasach przednaukowych, czy to w starożytnych wierzeniach, czy w rozwoju alchemii, kiedy próbowano znaleźć receptę na sztuczne życie, a przy okazji panaceum. Zostawiając jednak mity, wierzenia oraz szarlatanizm alchemii, dochodzimy do czasów nowożytnych, w których zaczęliśmy rozumieć, jak działa ludzki mózg. Wtedy to powstały pierwsze nowoczesne maszyny liczące oraz matematyczne podwaliny obliczeń komputerowych. Opisano też, że mózg działa w pewnym sensie jak skomplikowana sieć elektryczna zbudowana z neuronów, które odpowiadają (lub nie) na bodźce, a sygnały, które przesyła, można reprezentować w formie cyfrowej (komputerowej). Tak rozwinęły się programowalne komputery elektroniczne, z których pierwszym był Electronic Numerical Integrator and Computer (ENIAC). W latach 40. i 50. XX wieku wszystkie te osiągnięcia coraz częściej prowadziły do stawiania hipotez dotyczących tego, czy da się zaprogramo-



wać maszynę tak, żeby sama myślała. W 1950 roku Alan Turing w przełomowej pracy zastąpił formalne pytanie „Czy maszyny mogą myśleć?” pytaniem o ich zdolność do naśladowania inteligentnego zachowania poprzez posługiwanie się językiem naturalnym¹. W teście (nazwanym później testem Turinga) sędzia (człowiek) prowadzi pisemną konwersację (chat) z dwoma rozmówcami i ocenia, który z nich jest człowiekiem, a który maszyną. Jeśli nie jest w stanie tego określić, znaczy to, że maszyna zaliczyła test (czyli potrafi symulować inteligentne zachowanie człowieka w konwersacji). Ale czy zdolność odpowiadania na pytania zasługuje już na miano inteligencji? W 1955 roku inny prekursor AI John McCarthy wprowadził termin „sztuczna inteligencja” i zaproponował definicję, zgodnie z którą jest to „konstrukcja maszyny, o której działaniu

dałoby się powiedzieć, że jest podobna do ludzkich działań”. I znów trzeba było rozstrzygnąć, jakie działanie uznajemy za przejaw ludzkiej inteligencji.

Narodziny sztucznych neuronów

Zostawiając akademickie spory o definicję myślenia, część uczonych zabrała się do budowy czegoś, co naśladowałoby działanie ludzkiego mózgu. Współpraca neurobiologa Warrena McCullocha i logika Waltera Pittsa zaowocowała pracą, w której pojawił się koncept prostego elementu, tzn. cyfrowego neuronu². Cyfrowy neuron, nazwany później perceptronem, to podstawowy element współczesnych sieci neuronowych, który składa się z wejść (*inputs*), wag (*weights*) oraz wyjścia (*output*). Perceptron (podobnie jak każdy neuron) odbiera różne sygnały (informacje) za pomocą wejść (analogicznie jak dendryty reagujące na bodźce), wagi w perceptronie można porównać do siły sygnałów, która może być różna w zależności od połączeń synaptycznych, a wyjścia są analogią do synaps znajdujących się w zakończeniach aksonów. Perceptron „mnoży” sygnały wejściowe przez wagi, sumuje je, a potem, korzystając z mechanizmu aktywacji (*activation function*), decyduje, czy wynik tej sumy przekroczył próg decyzji. Jest to analogiczne do funkcji progów pobudzenia neuronu, który decyduje, czy przekaże sygnał dalej, czy nie – zgodnie z zasadą „wszystko albo nic”. Można by przytoczyć tutaj prosty przykład: perceptron analizuje 3 czynniki związane z meningokokowym zapaleniem opon mózgowo-rdzeniowych – pobyt w strefie endemicznego występowania choroby, gorączkę i wybroczyny widoczne na skórze. Każdy z tych objawów ma przypisaną wagę. Jeśli suma przekroczy ustalony próg, perceptron „diagnozuje” zapalenie opon. Takie podejście jest bardzo uproszczone, ale pozwala sobie wyobrazić, jak sztuczny neuron imituje podejmowanie decyzji, wykorzystując dane wejściowe oraz ich istotność. Neuron zaproponowany przez McCullocha i Pittsa oraz połączenia między takimi cyfrowymi neuronami można postrzegać jako prymitywną maszynę cyfrową, zdolną wykonywać podstawowe obliczenia logiki matematycznej (takie jak koniunkcja czy alternatywa). Praca ta, uznawana za fundamentalną dla dzisiejszych sieci neuronowych, miała

jednak dwa poważne ograniczenia: po pierwsze, neuron przyjmował dane i na nie odpowiadał, korzystając z wartości binarnych (0 lub 1), a jak nietrudno się domyślić, na świecie (a zwłaszcza w medycynie) niewiele jest rzeczy czarno-białych; po drugie, nie był zdolny do samodzielnego uczenia się, tylko trzeba było wpierw znaleźć rozwiązanie samemu, a potem taki neuron cyfrowy zaprogramować (czyli ustawić wyliczone wagi i próg decyzji). Cóż, mało inteligentnie zapowiadała się ta cała AI na początku, niemniej jednak Rubikon został przekroczony.

Pomimo tych ograniczeń powoli nadchodziła (pierwsza) era sztucznych sieci neuronowych. Już w 1957 roku Frank Rosenblatt razem ze swoim zespołem, zainspirowany wynikami pracy o cyfrowych neuronach McCullocha i Pittsa, przedstawił zarys Perceptronu, pierwszej fizycznej realizacji sieci neuronowej, którą ostatecznie zaprezentowano w 1960 roku^{3,4}. Alpha-Perceptron, bo tak nazwano pierwszą wersję sztucznej sieci neuronowej, był wyposażony w 400 czujników wejściowych (detektorów obrazu ułożonych rzędami 20 po 20), połączonych z 512 cyfrowymi neuronami (od tego momentu nazywanymi perceptronami), które z kolei łączyły się z systemem wyjścia odpowiedzialnym za reakcję na „problem” wykryty przez czujniki wejściowe. Zamiast korzystać z pojedynczych perceptronów, ułożono je w warstwę (*layer*) i zmuszono do współpracy, aby uzyskać końcową decyzję. We współczesnych sieciach neuronowych układa się, jedna po drugiej, bardzo dużo (setki) takich warstw (stąd też nazwa „głębokie/warstwowe sztuczne sieci neuronowe” [*deep neural networks*]).

Warstwa wejściowa przekształca dane wejściowe; stamtąd trafiają one do kolejnej warstwy, która je modyfikuje jeszcze bardziej, nadając im nowe cechy (*features*), i pozwala je zgłębiać na wielu poziomach abstrakcji aż do ostatniej warstwy, gdzie sztuczna sieć neuronowa podejmuje decyzje. Wczesne sieci neuronowe, często nazywane sieciami płytkimi (*shallow*), nie były w stanie znaleźć skomplikowanych (abstrakcyjnych) powiązań między danymi. Wracając do Alpha-Perceptronu, ta wyjątkowa maszyna, ważąca kilka ton, nie tylko mogła rozpoznawać proste symbole (takie jak niektóre litery alfabetu) wykrywane przez czujniki, ale – co najważniejsze – mogła się sama uczyć,

decydując o zmianie wag połączeń między 512 cyfrowymi neuronami ułożonymi w jedną warstwę! Samoucząca się maszyna, a tym samym AI, była już prawie na wyciągnięcie ręki!

Nadchodzi zima

Prezentacja tej „inteligentnej” maszyny wzbudziła zrozumiały zachwyty, ale komentarze samych twórców i finansowanie badań przez wojsko wywołały również liczne kontrowersje. Cytując za „The New York Times”: „Perceptron jest elektronicznym komputerem, od którego armia oczekuje, że będzie zdolny chodzić, mówić, widzieć, pisać, rozmnażać się (może bardziej precyzyjnie byłoby napisać «powielać się»?) i być świadomym swojej egzystencji”⁵. Obawy obawami, ale naukowcy zaczęli się prześcigać w przewidywaniach (dzieląc się raczej swoimi marzeniami i opiniami niż faktami i dowodami matematycznymi), za ile miesięcy lub lat AI rozwiąże wszystkie problemy ludzkości (uwzględniając te medyczne). Rozbudzone do granic (nie)możliwości oczekiwania (gdyby wtedy istniał TikTok, to pewnie pojawiłyby się hashtagi w stylu „#HypeNaAI”) przełożyły się na olbrzymie zainteresowanie badaniami nad AI (i tym samym na pieniądze), co skutkowało gwałtownym rozwojem specjalizujących się w tej dziedzinie jednostek naukowych (oczywiście tylko na zachód od Łaby).

Niestety pomimo gigantycznych inwestycji bardzo płytkie (jednowarstwowe) sztuczne sieci neuronowe nie potrafiły nie tylko się nauczyć rozwiązywania „trudnych problemów ludzkiej egzystencji”, ale nawet znaleźć prostego zastosowania przemysłowego, które gwarantowałyby szybki zwrot zainwestowanych pieniędzy. Możliwości obliczeniowe dostępnych w tamtych czasach komputerów i brak dużej ilości danych (a bardziej nawet brak możliwości technologicznych tworzenia dużych baz danych i ich przechowywania, bo warto pamiętać, że do komunikacji z komputerem wykorzystywano przede wszystkim papierowe taśmy dziurkowane), których to sieci neuronowe potrzebowały do uczenia się rozwiązania problemu, mocno hamowały rozwój sztucznych sieci neuronowych. Dodatkowo książka Minsky’ego i Paperta z 1969 roku, prezentująca matematyczne ograniczenia jednowarstwowych sztucznych sieci neuronowych (tzw. XOR – pro-



blem alternatywy wykluczającej), ostudziła zainteresowanie cyfrowymi neuronami i przyczyniła się do tzw. pierwszej zimy AI⁶.

Powstanie systemów eksperckich

Pomimo „pierwszej zimy sztucznych sieci neuronowych” AI rozwijała się nadal. Z powodu ograniczenia finansowania badań nad sieciami neuronowymi większość naukowców rozwijała inne metody AI, m.in. systemy, które obecnie nazywamy eksperckimi. Bazują one na dwóch podstawowych elementach, tj. wiedzy eksperckiej, którą należy zakodować w bazach danych, i mechanizmach decydowania o sposobach korzystania z tej wiedzy, najczęściej mających postać reguł logiki matematycznej. Reguły tworzone na podstawie specjalistycznej wiedzy w konkretnej dziedzinie i w bardzo podstawowej wersji można je sprowadzić choćby do przykładu: jeśli zachodzi pewna przesłanka (np. pacjent był w strefie endemicznego występowania choroby), wykonaj akcję (zmiierz pacjentowi temperaturę) albo sprawdź kolejny warunek (obejrzyj skórę w poszukiwaniu wybroczyn) itd., aż uzyskasz odpowiedź. Systemy eksperckie odnosiły pewne sukcesy w wąskich dziedzinach nauki i przemysłu, wymagały bowiem specjalistycznej infrastruktury oraz odpowiednio wykwalifikowanej kadry do budowy baz wiedzy (np. ekspertów me-

dycznych) i utrzymania takich systemów (dzisiaj byśmy ją nazwali działem IT).

Po raz pierwszy o AI w medycynie

Ale czy w czasie pierwszego zachwytu nad AI i pierwszych nią rozczarowań myślano o zastosowaniach w medycynie? Oczywiście! Na początku lat 70. XX wieku pojawiły się pierwsze inicjatywy naukowe próbujące łączyć środowisko medyczne z kształtującym się dopiero środowiskiem AI, a wśród nich powołany w 1973 roku The Research Resource on Computers in Biomedicine na Rutgers University czy the Stanford University Medical EXperimental computer for Artificial Intelligence in Medicine (SUMEX-AIM). Już w 1975 roku odbyła się pierwsza konferencja poświęcona AI w medycynie⁷. Pośród wielu przykładów jej zastosowań warto przytoczyć kilka. Pierwszy to system ekspercki DIALOG (z czasem nazwany INTERNIST [!]) – model, który łączył dane z wywiadu (objawy), wyniki badań laboratoryjnych oraz historię pacjenta do tworzenia rankingów potencjalnych chorób dostępnych w bazie danych systemu^{8,9}. Mimo że system ten budził wiele nadziei szczególnie tam, gdzie brakowało wykwalifikowanych internistów, nie wyszedł poza etap projektu naukowego i nie przyjął się w praktyce lekarskiej. I nie było to spowodowane tym, że bazy danych uzupełniali studenci medycyny, wprowadzając godzinami do systemu wyniki badań klinicznych publikowanych w recenzowanych czasopiśmie naukowych. Główne problemy wynikały z trudności i czasochłonności obsługi (konsultacja komputerowa trwała 60–90 min), konieczności posiadania dostępu do dużego komputera (czasy komputerów osobistych miały dopiero nadejść) oraz zorganizowania szkolenia w zakresie korzystania z komputera z zainstalowanym systemem INTERNIST¹⁰. Z perspektywy medycznej z kolei największą przeszkodą była zapewne ograniczona zdolność diagnozowania skomplikowanych przypadków (tzn. pacjentów z wielochorobowością).

Innymi modelami zaprezentowanymi na wspomnianej konferencji były CASNet, komputerowy system diagnozowania i rekomendacji leczenia jaskry¹¹, oraz MYCIN, który miał identyfikować chorobę bakteryjną i rekomendować antybiotyk oraz

jego dawkę dostosowaną do masy ciała chorego¹². W systemie eksperckim MYCIN chorobę bakteryjną identyfikowano na podstawie odpowiedzi „tak” lub „nie” na zaprogramowane pytania (np. czy bakteria jest Gram-dodatnia, czy ma pałeczkowaty kształt). Każdej odpowiedzi przypisywano sugerowane rozpoznanie i jego prawdopodobieństwo (np. zakażenie *Pseudomonas aeruginosa*, 0,4). Żadnemu z powyższych systemów eksperckich nie udało się przekroczyć progu akceptacji, żeby zostać narzędziem w rękach lekarzy. Podobnie jak w przypadku INTERNIST, czasochłonność i trudności w korzystaniu z MYCIN (np. konieczność ręcznego wpisywania wszystkich danych pacjenta) oraz bariery technologiczne (np. dotyczące dostępności komputerów) mocno ograniczały potencjalne zastosowania tego systemu w codziennej praktyce lekarskiej, mimo że MYCIN w ślepym teście wykazał się podobną dokładnością co specjaliści w zakresie leczenia zapalenia opon mózgowo-rdzeniowych¹³.

Warto też wspomnieć, że żaden z powyższych systemów nie korzystał ze sztucznych sieci neuronowych, tylko z przeżywających swój złoty wiek systemów eksperckich, działających na podstawie ustrukturyzowanych reguł (ok. 200 w modelu MYCIN) opracowanych dzięki współpracy ze specjalistami. Sztuczna inteligencja oparta na systemach eksperckich, takich jak MYCIN czy INTERNIST, udowodniła, że może osiągać obiecujące wyniki w wąskiej specjalizacji. A gdybyśmy tylko mogli zaprogramować większy (niekończony?) zasób reguł, to pewnie takie systemy już dawno rozwiązałyby wszystkie problemy, nie tylko medyczne...

Sztuczne sieci neuronowe – powoli naprzód i w górę!

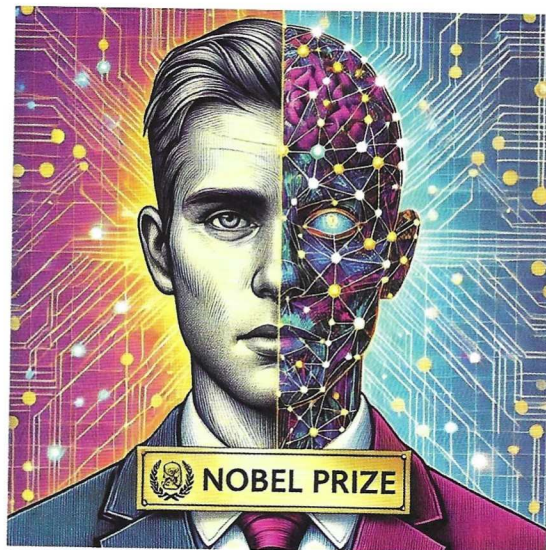
Jednocześnie, pomimo braku finansowania, w zaciszu uniwersyteckim próbowano zbudować solidne teoretyczne fundamenty sztucznych sieci neuronowych oraz znaleźć praktyczne rozwiązania problemów, które trapiły twórców ich prototypów.

W latach 80. XX wieku fizyk John Hopfield zaproponował model jednowarstwowej sieci neuronowej (nazywanej później siecią Hopfielda), w której każdy cyfrowy neuron był połączony ze wszystki-

mi innymi neuronami (z wyjątkiem samego siebie)¹⁴. Taka struktura pozwalała na rekurencyjne uczenie się i późniejsze rozpoznawanie wzorców na podstawie skojarzeń z innym wzorcem albo nawet jego częścią. Działanie sieci Hopfielda można porównać do pamięci asocjacyjnej, która umożliwia sieci „przyciąganie” nowego wzorca do najbliższego podobnego wzorca w pamięci (przez analogie, np. „kiedyś już miałem podobny przypadek, kiedy przyszedł do mnie pacjent z temperaturą powyżej 38,0°C, który wrócił ze strefy endemicznego występowania choroby”). Dodatkowo Hopfield zaproponował koncepcję minimalizacji energii takiej sieci, dążąc do jej stabilności (czyli poszukiwania stanu, w którym sieć może efektywnie odpowiadać na postawione pytania), na wzór systemów znanych z biologii czy fizyki, które także naturalnie dążą do minimalizacji energii.

W tym samym czasie Geoffrey Hinton ze współpracownikami rozważał teoretyczne i praktyczne możliwości wielowarstwowych sieci neuronowych, a także pracował nad tzw. maszynami Boltzmann¹⁵. Maszyna Boltzmann to także sztuczna sieć neuronowa podobna do powyżej opisanej sieci Hopfielda, tyle że wyposażona w element stochastyczny (losowy). Wprowadzenie elementu losowości i wielowarstwowości pozwala sztucznej sieci neuronowej na poszukiwanie bardziej złożonych wzorców w danych. W 2024 roku obaj naukowcy, John Hopfield i Geoffrey Hinton, zostali uhonorowani Nagrodą Nobla w dziedzinie fizyki za ich wkład w fundamentalne odkrycia i wynalazki, które umożliwiły uczenie maszynowe z wykorzystaniem sztucznych sieci neuronowych. Ich prace stanowią fundament współczesnej AI. Model pamięci asocjacyjnej Hopfielda i stochastyczne podejście Hintona zainspirowały rozwój głębokich sieci neuronowych, które 30 lat później zrewolucjonizowały AI i znalazły zastosowanie w licznych dziedzinach, w tym w medycynie.

Rozwiązanie problemu, jak uczyć sztuczne sieci neuronowe mające więcej niż jedną warstwę (czyli więcej niż Alpha-Perceptron w 1957 r.), zaproponował Paul Werbos, opracowując algorytm propagacji wstecznej błędów (*back-propagation*)¹⁶. Często jako prekursorów tego rozwiązania podaje się także innych naukowców^{17,18}. Algorytm propagacji wstecznej uaktualnia wszystkie wagi, korzystając



z informacji o tym, czy sieć neuronowa udzieliła poprawnej czy błędnej odpowiedzi, na podstawie porównania tej odpowiedzi z odpowiedzią prawidłową. Informacja o błędzie jest przenoszona z warstwy wyjściowej do warstwy wejściowej sztucznej sieci neuronowej (czyli wstecz, stąd nazwa tego algorytmu – prawda, że sztuczne sieci neuronowe są proste?): neuron po neuronie, warstwa po warstwie. Propagowanie błędu pozwala (za pomocą reguły łańcuchowej znanej wszystkim, którzy uczęszczali na lekcje matematyki w szkole średniej; jeśli nie, to warto sobie przypomnieć emocje towarzyszące pytaniu: „No, młody człowieku, jak policzyć pochodną funkcji złożonej?”) na obliczenie, o ile powinna się zmienić waga perceptronu, żeby zmniejszyć błąd popełniany przez sztuczna sieć neuronową. Proces uaktualniania wag sieci neuronowej, czyli uczenia się albo trenowania sztucznej sieci neuronowej, powtarza się tak długo, aż sieć zacznie osiągać zadowalające wyniki dla zebranych danych. Badaniami nad algorytmem propagacji wstecznej zajmował się też wspomniany wcześniej Hinton, prezentując wyniki sugerujące, że głębokie sieci neuronowe za pomocą tego algorytmu mogą się uczyć bardziej abstrakcyjnych właściwości danych¹⁹.

Sieci splotowe i poszukiwanie abstrakcji

Jeszcze innym przełomowym wynalazkiem, który pojawił się w tamtym czasie, były splotowe (konwolucyjne) sieci neuronowe (CNN). Sieci te potrafiły nie tylko – jak ich wielowarstwowi poprzednicy – uczyć się udzielania poprawnej odpowiedzi na podstawie dostępnych danych, ale również – co szczególnie ważne – samodzielnie się uczyć, jakie cechy (*features*) danych przestrzennych (takich jak zdjęcia lub dźwięk) są potrzebne do jej udzielenia! Współczesne głębokie CNN są zdolne do samodzielnego uczenia się skomplikowanych wzorców i zależności między danymi, nawet jeśli nikt im nie podpowie, na jaką cechę mają zwrócić uwagę. Na przykład głęboka CNN, która rozpoznaje, gdzie na zdjęciu znajduje się twarz (w taką funkcję jest wyposażony prawie każdy smartfon dostępny obecnie na rynku), w niższych warstwach uczy się podstawowych cech, takich jak krawędzie lub tekstury, zaś w warstwach środkowych łączy podstawowe cechy w bardziej złożone kształty (kilka różnych krawędzi może „budować” fragmenty oczu lub nosa), natomiast warstwy głębokie kojarzą wszystkie wcześniejsze cechy i umożliwiają sieci rozpoznanie twarzy. W radiologii (np. mammografii) kolejne warstwy CNN rozpoznają coraz bardziej złożone cechy. Niższe warstwy, podobnie jak w przypadku zdjęć twarzy, identyfikują podstawowe krawędzie i tekstury, na przykład żeby odróżnić tkankę miękką (ciemne przestrzenie) od mikrozwapnień (jasne punkty). Środkowe warstwy wychwytyją bardziej złożone cechy, takie jak skupiska jasnych i ciemnych obszarów i ich rozłożenie (czy znajdują się w grupach albo w liniach). Warstwy głębokie natomiast poszukują abstrakcyjnych cech charakteryzujących większy obszar tkanki, co pozwala na ocenę otoczenia mikrozwapnień (kontekst występowania) oraz ich umiejscowienie w piersi.

Sam pomyśl, że każdy wzorzec składa się z podstawowych elementów, nie był tak zupełnie nowy. W 1957 roku David Hubel i Torsten Wiesel odkryli (korzystając z pomocy mikroelektrod i kota), że wyświetlanie różnych wzorców na ekranie, na który spogląda mały kotek, powoduje wzbudzenie różnych neuronów w jego korze wzrokowej^{20,21}.

Postulowali też, że rozpoznawanie wzorców jest hierarchiczne, tzn. komórki niższego rzędu identyfikują proste, a komórki wyższego rzędu bardziej skomplikowane kształty, na przykład litery czy cyfry. Ponad ćwierć wieku współpracy tych naukowców zaowocowało w 1981 roku Nagrodą Nobla w dziedzinie fizjologii lub medycyny za przełomowe odkrycia dotyczące przetwarzania informacji wzrokowej w mózgu²². Ich badania zrewolucjonizowały rozumienie, jak mózg interpretuje bodźce wzrokowe, i miały ogromny wpływ na rozwój neuronauki. W pewien sposób były zatem inspiracją dla twórców głębokich sieci neuronowych, takich jak CNN.

Przedwiośnie zastosowań głębokich sieci neuronowych

Budowanie głębokich CNN otworzyło drzwi do przełomowych zastosowań. Już w 1980 roku Kunihiko Fukushima zaproponował model (*neocognition*), który potrafił rozpoznawać wzorce (np. cyfry), ucząc się, że składają się one z pewnych kształtów geometrycznych i podobieństw między nimi²³. Sama koncepcja rozpoznawania wzorców na podstawie agregacji coraz bardziej skomplikowanych cech dzięki budowaniu coraz głębszych sieci neuronowych zaczęła przyciągać uwagę kolejnych naukowców. Podążając tym tropem, rozwijano CNN, które zyskały zdolność rozpoznawania ręcznie pisanych cyfr, dzięki czemu mogły automatycznie odczytywać kody pocztowe z listów²⁴. Pojawiły się też pierwsze zastosowania CNN w medycynie, takie jak analiza mammograficzna w celu wykrywania mikrozwapnień i wczesnych oznak raka piersi^{25,26}. Splotowe sieci neuronowe wykorzystywano do interpretacji zapisów elektrokardiograficznych w celu identyfikacji arytmii oraz innych zaburzeń rytmu serca. Wczesne systemy potrafiły klasyfikować wzorce zapisów elektrokardiograficznych i wykrywać potencjalnie groźne stany, takie jak migotanie przedsionków, co mogło zapobiegać nagłym zdarzeniom sercowo-naczyniowym²⁷. Tak oto, mimo że w latach 80. ubiegłego wieku dominowały systemy eksperckie, postęp w badaniach nad sztucznymi sieciami neuronowymi powoli budził kolejną wiosnę. Niemniej jednak ich stosowanie w medycynie

(i nie tylko) nadal wiązało się z poważnymi ograniczeniami. Bariera dostępu do komputera powoli zanikała (przynajmniej na zachód od Łaby) dzięki szybkiemu rozwojowi komputerów osobistych. Problemem pozostawał brak wystarczająco dużej mocy obliczeniowej komputerów, co ograniczało możliwość uczenia się na dużych zbiorach danych. Przykładowo, wytrenowanie trzywarstwowej sieci neuronowej (czyli bardzo płytkiej jak na dzisiejsze czasy) do rozpoznawania znaków z kodu pocztowego na listach (baza danych stworzona do tego eksperymentu gromadziła ok. 10 000 zdjęć znaków kodu pocztowego, a każde zdjęcie miało rozmiar 16 na 16 pikseli [razem 256 pikseli]), co jest niczym w porównaniu ze współczesnym zdjęciem zrobionym najprostszym cyfrowym aparatem fotograficznym ukrytym w zwykłym telefonie, które może mieć więcej niż 50 mln pikseli) zajmowało około 3 dni – i to mimo korzystania ze specjalnego komputera wysokiej wydajności (*high-performance workstation*)²⁴. Kolejnym problemem był dostęp do danych medycznych. Nie miały one jeszcze formy cyfrowej, ponieważ standard ich zapisu, przechowywania i przesyłania (DICOM) miał się dopiero pojawić w 1993 roku. Dodatkowo decyzje podejmowane przez sztuczne sieci neuronowe charakteryzowały się – w odróżnieniu od systemów eksperckich – brakiem interpretowalności lub wyjaśnialności. Sztuczne sieci neuronowe często były postrzegane jako „czarne skrzynki” (*black box*) z tego powodu, że lekarzom trudno było zrozumieć, na jakiej podstawie system ustalił rozpoznanie. Większe nadzieje nadal pokładano zatem w systemach eksperckich, a także w rodzącej się w tym samym czasie AI opartej na modelowaniu zachowań (*behavioral artificial intelligence*)²⁸.

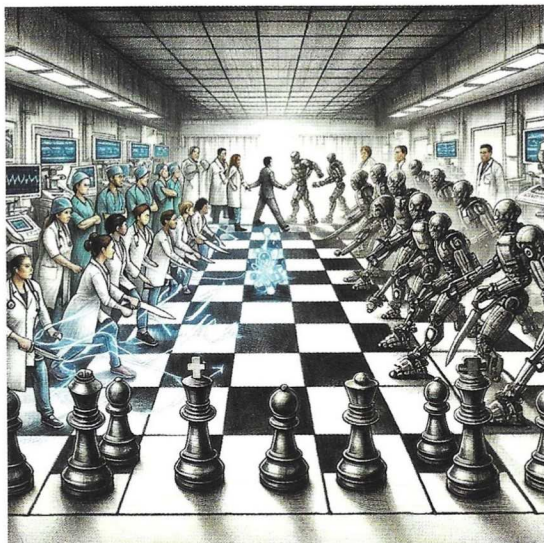
A może by tak porzucić sieci neuronowe i modelować zachowania?

Brooks zaproponował alternatywne podejście do AI oparte na modelowaniu bezpośredniej interakcji z otoczeniem, w której to koncepcji roboty nie miały abstrakcyjnej reprezentacji danych (jak w sztucznych sieciach neuronowych) ani zasad logiki ich przetwarzania (jak w systemach eksperckich). Roboty wyposażone w behawioralną AI

miały natomiast reagować na bodźce zewnętrzne i podejmować decyzje na podstawie prostych reguł zachowania²⁸. Reguły zachowania z kolei miały podlegać zasadzie subsumpcji, która organizowała interakcje hierarchicznie: zadanie wyższego poziomu (o wyższym priorytecie, np. wykonaj resuscytację krążeniowo-oddechową u pacjenta bez tętna i oddechu) mogło „subsumować” zadanie niższego poziomu (np. wytłumacz nowo przyjętej stażystce lub stażystce, jaką lubisz kawę). Połączenie hierarchii ważności i prostych modeli reakcji na bodźce zewnętrzne ułatwiało dostosowanie i skalowanie AI do różnych zadań i środowisk (kawę pije się nie tylko w szpitalu). Kiedy zatem patrzymy na inteligentny odkurzacz (gdybym taki miał, nie musiałbym sam odkurzać) – który wydaje się naprawdę inteligentny, bo trafiając na ścianę (lub nogę mojej wspaniałej żony albo kota mojej córki), zmienia kierunek sprzątania, a gdy ma pełny zbiornik kurzu, przypomina (do skutku) o potrzebie jego opróżnienia – możemy śmiało stwierdzić, że prace Brooksa i jego następców wywarły kolosalny wpływ na rozwój robotyki i inteligentnych systemów działających w dynamicznych środowiskach (gdzie decyzje trzeba podejmować natychmiastowo). Choć sam Brooks nie tworzył systemów AI dla medycyny, podejście oparte na modelowaniu reakcji w następstwie obserwowanych objawów bardzo mocno się łączy z opieką zdrowotną. Zamiast skomplikowanych i trudnych do wdrożenia systemów podejście behawioralne proponuje prosty sposób modelowania interakcji technologii z pacjentem i środowiskiem opieki zdrowotnej, szczególnie w stanach nagłych.

Człowiek kontra maszyna – pierwsza krew

Kolejny okres rozwoju AI przerywały od czasu do czasu doniesienia prasowe, takie jak wygrana (w 1996 r.) i przegrana (1997 r.) człowieka w pojedynku szachowym między mistrzem Garrim Kasparowem a komputerem IBM Deep Blue²⁹. Deep Blue był zaawansowanym systemem eksperckim do gry w szachy, zbudowanym na bazie wiedzy i doświadczeń innych mistrzów szachowych i wykorzystującym potencjał obliczeniowy superkomputerów (Deep Blue miał możliwość



wykonania 200 mln operacji szachowych na sekundę) do sprawdzenia każdego możliwego ruchu szachowego. Choć Deep Blue nie korzystał ze sztucznych sieci neuronowych, zwycięstwo nad mistrzem szachowym w 1997 roku wielu uznaje za symboliczną datę pierwszego intelektualnego zwycięstwa maszyny nad człowiekiem. Choć cóż to za inteligencja sprawdzać każde możliwe rozwiązanie i nic się z niego nie uczyć? Innym razem prasa donosiła o konkursie organizowanym przez Defense Advanced Research Projects Agency (DARPA) dla autonomicznych samochodów zdolnych przejechać 132 mile gorącej pustyni Nevady³⁰. Warto odnotować, że w pierwszej edycji żaden pojazd nie pokonał dystansu większego niż 7,5 mili. Rok później inteligentny samochód zaprojektowany przez zespół z Uniwersytetu Stanforda dotarł do mety w niespełna 7 godzin. Mniej medialnym wydarzeniem, które miało duży wpływ na rozwój obecnej AI, było skonstruowanie i wprowadzenie na rynek kart graficznych (GPU) przez firmę NVIDIA w 1999 roku. Na początku GPU cieszyły się dużym powodzeniem wśród graczy komputerowych i deweloperów aplikacji multimedialnych, ale ich konstrukcja oparta na dużej liczbie równoległych jednostek obliczeniowych miała się już niebawem okazać doskonałym panaceum na jeden z problemów, z którym borykały się sztuczne sieci neuronowe od swoich narodzin, czyli ograniczenia obliczeniowe. Mimo że pierwsze GPU projektowa-

no z myślą o optymalizacji grafiki, można je było z powodzeniem wykorzystywać do intensywnych obliczeń (głębokie sieci neuronowe to duża ilość mnożeń i dodawań), kluczowych w rozwijaniu AI, która potrafi się sama uczyć. Brakowało jeszcze zaledwie jednego elementu... danych!

O kotach, psach i innych rzeczach, które (na pierwszy rzut oka) nijak się mają do medycyny

Jak zdobyć dużo danych, żeby trenować bardzo głębokie sieci neuronowe? Najlepiej ściągnąć je z Internetu! Na taki pomysł wpadli naukowcy z Princeton pod kierownictwem Fei-Fei Li i rozpoczęli prace nad ImageNet, największą bazą zdjęć z opisem ich zawartości³¹. Baza danych osiągnęła rozmiar 14 milionów zdjęć z hierarchicznym opisem (zwierzę → ssak → pies → owczarek podhalański), a samo opisywanie zdjęć zajęło 2,5 roku. Mając do dyspozycji masowy zbiór danych obrazowych odpowiednich do trenowania głębokich sieci spłotowych (bo zdjęcia to dane przestrzenne), rozpisano konkurs na rozpoznawanie obiektów znajdujących się na zdjęciach – ImageNet Large Scale Visual Recognition Challenge (ILSVRC), który odbywał się corocznie od 2010 do 2017 roku³². Na spektakularne wyniki nie trzeba było długo czekać. W 2012 roku zespół pod kierunkiem Geoffreya Hinton (to ten przywołany już wcześniej pan od Nobla z fizyki w 2024 r.) oraz Alexa Krizhevsky'ego i Ilyi Sutskevera przedstawił model AlexNet, pierwszą głęboką (pięciowarstwową) CNN, która wygrała konkurs ILSVRC z olbrzymią przewagą, osiągając o 10% mniejszy błąd niż konkurencja³³. AlexNet był jednym z pierwszych modeli, który używał GPU do przyspieszenia trenowania sieci, co pozwoliło na trenowanie głębszej CNN. Sukces AlexNet pokazał, że głębokie CNN mogą znacząco przewyższać tradycyjne podejścia do rozpoznawania obrazów, i zainspirował środowisko naukowe do rozpoczęcia na nowo badań nad głębokim uczeniem³⁴. W kolejnych latach konkurs ILSVRC tchnął nowe życie w rozwój coraz bardziej złożonych (głębokich) i wydajnych architektur sieci, takich jak 19-warstwowy VGG (VGG2014), 22-warstwowa GoogLeNet³⁵ (później przemianowana na Inception w nawiązaniu do cytatu „Mu-

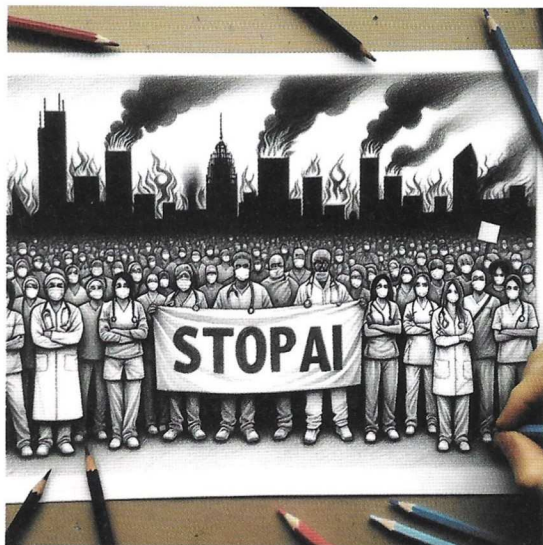
simy zejść głębiej” z filmu „Incepcja” w reżyserii Christophera Nolana) oraz ponad 100-warstwowy ResNet³⁶, które były kolejne rekordy dokładności w klasyfikacji obrazów, w 2015 roku osiągając już poziom zbliżony do człowieka. ImageNet oraz konkurs ILSVRC przyciągnęły uwagę badaczy z różnych dziedzin, napędzając rozwój głębokiego uczenia nie tylko w rozpoznawaniu obrazów, ale także w przetwarzaniu języka naturalnego, rozpoznawaniu mowy i innych obszarach AI. Kolejne spektakularne sukcesy w stosowaniu głębokich sieci neuronowych zmotywowały także rządowe agencje naukowe (choć nie w każdym państwie) i przemysł do inwestowania w badania nad głębokim uczeniem oraz rozwój infrastruktury obliczeniowej (takiej jak GPU), co doprowadziło do lawinowego zwiększenia liczby publikacji z dziedziny AI – podwaja się ona co 23 miesiące³⁷. ImageNet pomógł wielu firmom wprowadzić AI do swoich produktów i usług w autonomicznych samochodach, robotach, systemach rozpoznawania twarzy, aplikacjach medycznych i wielu innych. ImageNet stał się fundamentem współczesnej AI i symbolem jej ogromnego potencjału. A jak w tym błyskawicznym rozwoju AI odnalazła się medycyna?

Czy leci z nami radiolog?

Postęp w zakresie uczenia głębokiego i doświadczenia z ImageNet miały ogromny wpływ na rozwój technologii AI w zastosowaniach medycznych, szczególnie w analizie obrazów diagnostycznych, takich jak radiogram czy obrazy uzyskane w tomografii komputerowej lub obrazowania metodą rezonansu magnetycznego, a także otworzyły potencjał do automatyzacji wykrywania, klasyfikacji i opisywania zmian chorobowych z precyzją zbliżoną do osiąganą przez doświadczonych radiologów, a czasem nawet większą. Pierwszym przykładem zastosowania głębokich CNN w medycynie była sieć wspomagająca analizę zmian skórnych (takich jak znamiona), która umożliwiała szybsze i bardziej dostępne rozpoznawanie potencjalnie groźnych zmian (np. czerniaka) na podstawie zdjęć skóry³⁸. Wspomniana wcześniej CNN Inception została wytrenowana na ponad 127 tysiącach zdjęć zmian skórnych, po czym jakość udzielanych przez nią odpowiedzi porównano z jakością odpowiedzi

21 certyfikowanych dermatologów. Z racji tego, że w nauce, a w czasopiśmie „Nature” szczególnie, rzadko się omawia negatywne wyniki, nietrudno przewidzieć, jak wypadło starcie człowieka z maszyną... Przedstawiona CNN działała nie gorzej niż specjaliści, autorzy postulowali zatem, że narzędzie, które będzie z niej korzystać, może mieć duże znaczenie kliniczne, wpływając m.in. na rozszerzenie zakresu kompetencji lekarzy podstawowej opieki zdrowotnej oraz wspomaganie procesu podejmowania decyzji klinicznych przez specjalistów dermatologii. Choć diagnoza dermatologiczna rzadko się sprowadza wyłącznie do obejrzenia zmian skórnych, pierwszy krok na drodze do wprowadzania sztucznych sieci neuronowych, dorównujących dokładnością specjalistom, został postawiony. Następne miesiące przynosiły nowe raporty naukowe o kolejnych sztucznych sieciach neuronowych, które działały równie dobrze jak specjaliści (a czasem nawet od nich lepiej).

Innym zadaniem, w którym sztuczne sieci neuronowe sprawdziły się nie gorzej niż specjaliści, było ustalanie potencjalnego ryzyka ostrej i przewlekłej utraty wzroku i kierowanie pacjentów wymagających szybkiej, pozaplanowej oceny tego stanu do specjalisty na podstawie wyniku optycznej tomografii koherencyjnej³⁹. Trening oparty na trójwymiarowych obrazach tomograficznych okazał się możliwy i tym samym pokonano kolejną barierę technologiczną, co otworzyło perspektywę różnorodnego zastosowania AI w radiologii. Wraz z rozwojem sztucznych sieci neuronowych w medycynie zaczęto wysuwać postulat powszechnego dostępu do danych medycznych (np. MIMIC-III był bazą danych ze szpitala specjalistycznego, zawierającą informacje dotyczące pacjentów przyjętych na oddziały intensywnej terapii). Dostępne dane, które miały m.in. wspierać badania naukowe, obejmowały parametry życiowe, przyjmowane leki, wyniki badań laboratoryjnych, wnioski z obserwacji i notatki sporządzane przez personel medyczny, bilans płynów, kody procedur, kody diagnostyczne, raporty z badań obrazowych, długość pobytu w szpitalu oraz dane dotyczące przeżycia^{40,41}. Otwarcie dostępu do danych rozbudziło zainteresowanie – zarówno naukowców, jak i przemysłu – zastosowaniami AI w medycynie. Pojawiły się więc raporty naukowe, które przed-



stawiały, jak zbudować głębokie sieci neuronowe rozpoznające zapalenie płuc⁴² lub kardiomiopatię przerostową⁴³ na podstawie radiogramów klatki piersiowej. Wspominany już wcześniej kilka razy Geoffrey Hinton, jeden z ojców głębokich sieci neuronowych, w czasie konferencji w 2016 roku zszokował środowisko radiologiczne, mówiąc: „Powinniśmy przestać szkolić radiologów już teraz. To całkowicie oczywiste, że w ciągu 5 lat uczenie głębokie będzie osiągać lepsze wyniki niż oni”⁴⁴. Z jasnych powodów ten pozornie banalny komentarz wywołał głęboki odzew. Co miały sobie pomyśleć osoby rozważające specjalizację w tej dziedzinie? Oczywiście tylko jedno – że AI może uczynić ich zawód przestarzałym, zanim skończą specjalizację. Obawy przed utratą perspektyw zawodowych często popychały takie osoby na inne ścieżki specjalizacyjne. Nieco wcześniej Elon Musk przewidywał, że jesteśmy bliżej samochodów autonomicznych, niż ktokolwiek myśli. „Uważam – oświadczył – że jesteśmy mniej niż 2 lata od pełnej autonomii, bezpieczniejszej niż jazda ludzka, ale przepisy będą wymagały przynajmniej kolejnego roku”⁴⁵. Sprawy AI istotnie nabierają tempa i już w 2017 roku amerykańska Agencja ds. Żywności i Leków (Food and Drug Administration – FDA) zaakceptowała pierwsze rozwiązanie wykorzystujące AI do użycia w szpitalach. Mały postęp technologiczny, ale wielki krok ludzkości w kierunku „nowej medycyny”⁴⁶.

Pandemia

Parafrazując innego noblistę, „2019 był to dziwny rok, w którym rozmaite znaki na niebie i ziemi zwiastowały jakoweś klęski i nadzwyczajne zdarzenia”. Zamiast zamieszek wzniesionych przez bezrobotnych radiologów albo kolejek po autonomiczne samochody mieliśmy pandemię COVID-19. Gigantyczna maszyna ratowania ludzkości ruszyła bardzo szybko. Badanie RECOVERY zyskało aprobatę i w kilka dni zrekrutowano pierwszego uczestnika, a w ciągu 3 miesięcy naukowcy dostarczyli pierwsze sprawdzone skuteczne leczenie⁴⁷. Podobną szybkość działania obserwowaliśmy w badaniach nad skuteczną szczepionką. A gdzie była AI w walce z pandemią? Skoro korzystaliśmy z niej do interpretacji objawów zapalenia płuc, wydawało się oczywiste, że powinna mieć duży potencjał w szybkim i dokładnym wykrywaniu oraz prognozowaniu COVID-19 na podstawie standardowych radiogramów klatki piersiowej lub obrazów tomografii komputerowej. Pomimo ogromnego nakładu pracy w celu opracowania modeli AI do diagnostyki i prognozowania COVID-19 w czasie pandemii żaden nie nadawał się do klinicznego zastosowania⁴⁸. Dlaczego? Raport Instytutu Alana Turinga odpowiada na to pytanie wprost: brak danych. Sztuczną inteligencję możemy tworzyć tylko, korzystając z danych – dużej liczby danych, a tymczasem w czasie pandemii ujawniły się wszystkie bolączki systemów obsługujących dane medyczne: problemy z dostępnością i standaryzacją, a także nierówności w dostępie badaczy do danych⁴⁹. Wniosek na przyszłość jest jasny: lepsze dane pozwoliłyby na lepszą reakcję.

Podsumowanie

Obecnie dysponujemy setkami narzędzi AI zatwierdzonych przez FDA do pracy w warunkach klinicznych. Sztuczna inteligencja jako narzędzie wspierające wykrywanie raka piersi w programie przesiewowym opartym na podwójnym odczycie mammografii okazała się lepsza lub nie gorsza od ludzkiego podwójnego odczytu we wszystkich wskaźnikach, a korzystanie z niej pozwoliło ograniczyć liczbę przypadków wymagających drugiego odczytu przez człowieka nawet o 87%⁵⁰. Zastosowanie

wania AI obejmują przede wszystkim radiologię, ale nie tylko, szczególnie teraz, w czasach wielkich modeli językowych.

Na zakończenie wróć do pytania, które usłyszałem od starszego lekarza w czasie konferencji. „Czy za 5 lat AI zastąpi nas w przychodni?”. Odpowiem, że nie wiem. Wspomnę tylko, że ledwie chwilę temu, kiedy już prawie skończyłem pisać ten tekst, zauważyłem przez okno mojego sąsiada, specjalistę radiologii. Otworzył drzwi samochodu po stronie kierowcy, odpalił silnik i pewnie już dojeżdża do szpitala. A zatem ewolucja...

PIŚMIENNICTWO

- Turing A.M.: Computing machinery and intelligence. *Mind*, 1950; 59: 433–460
- McCulloch W.S., Pitts W.: A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, 1943; 5: 115–133
- Rosenblatt F.: The perception: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 1957; 65: 386–408
- Rosenblatt F.: Principles of neurodynamics: perceptrons and the theory of brain mechanisms. Buffalo (NY), Cornell University, 1962
- New navy device learns by doing. Psychologist shows embryo of computer designed to read and grow wiser. *New York Times*, 1958. www.nytimes.com/1958/07/08/archives/new-navy-device-learns-by-doing-psychologist-shows-embryo-of.html (dostęp: 25.11.2024)
- Minsky M., Papert S.A.: Perceptrons, reissue of the 1988 expanded edition with a new foreword by Léon Bottou: an introduction to computational geometry. MIT Press, 2017
- Kulikowski C.A.: An opening chapter of the first generation of artificial intelligence in medicine: the first rutgers AIM workshop, June 1975. *Yearb. Med. Inform.*, 2015; 10: 227–233
- Pople H.E., Myers J.D., Miller R.A.: DIALOG: a model of diagnostic logic for internal medicine. *IJCAI*, 1975. www.ijcai.org/Proceedings/75/Papers/127.pdf (dostęp: 25.11.2024)
- Pople H.E. Jr.: Presentation of the INTERNIST System Proc 2nd AIM Workshop. New Brunswick (NY), Rutgers University, 1976
- Miller R.A., Pople H.E. Jr, Myers J.D.: Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. Computer-assisted medical decision making. New York (NY), Springer New York, 1985: 139–158
- Weiss S., Kulikowski C.A., Safir A.: Glaucoma consultation by computer. *Comput. Biol. Med.*, 1978; 8: 25–40
- Shortliffe E.H., Davis R., Axline S.G. i wsp.: Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. *Comput. Biomed. Res.*, 1975; 8: 303–320
- Victor L.Y., Fagan L.M., Wraith S.M. i wsp.: Antimicrobial selection by a computer: a blinded evaluation by infectious diseases experts. *JAMA*, 1979; 242: 1279–1282
- Hopfield J.J.: Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci.*, 1982; 79: 2554–2558
- Ackley D.H., Hinton G.E., Sejnowski T.J.: A learning algorithm for Boltzmann machines. *Cog. Sci.*, 1985; 9: 147–169
- Werbos P.J.: Applications of advances in nonlinear sensitivity analysis. System modeling and optimization: Proceedings of the 10th IFIP Conference New York City, USA, August 31–September 4, 1981. Berlin–Heidelberg, Springer, 2005
- Amari S.I.: A theory of adaptive pattern classifier. *IEEE Trans.*, 1967; EC-16: 279–307 (w j. jap. 1965)
- Linnainmaa S.: The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors (Masters). University of Helsinki, 1970: 6–7
- Rumelhart D.E., Hinton G.E., Williams R.J.: Learning representations by back-propagating errors. *Nature*, 1986; 323: 533–536
- Hubel D.H., Wiesel T.N.: Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.*, 1959; 148: 574–591
- Hubel D.H., Wiesel T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 1962; 160: 106–154
- Hubel D.H., Wiesel T.N.: Brain and visual perception: the story of a 25-year collaboration. Oxford University Press, 2004
- Fukushima K.: Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.*, 1980; 36: 193–202
- LeCun Y., Boser B., Denker J.S. i wsp.: Backpropagation applied to handwritten zip code recognition. *Neural. Comput.*, 1989; 1: 541–551
- Wu Y., Doi K., Giger M.L., Nishikawa R.M.: Computerized detection of clustered microcalcifications in digital mammograms: applications of artificial neural networks. *Med. Phys.*, 1992; 19: 555–560
- Zhang W., Doi K., Giger M.L. i wsp.: Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. *Med. Phys.*, 1994; 21: 517–524
- Baxt W.G.: Application of artificial neural networks to clinical medicine. *Lancet*, 1995; 346: 1135–1138
- Brooks R.A.: New approaches to robotics. *Science*, 1991; 253: 1227–1232
- Deep Blue. 1996. www.ibm.com/history/deep-blue (dostęp: 25.11.2024)
- The Grand Challenge. DARPA 2004. www.darpa.mil/about-us/timeline-grand-challenge-for-autonomous-vehicles (dostęp: 25.11.2024)
- Deng J., Dong W., Socher L.J. i wsp.: Imagenet: a large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition. *Ieee*, 2009: 248–255
- Russakovsky O., Deng J., Su H. i wsp.: Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.*, 2015; 115: 211–252
- Krizhevsky A., Sutskever I., Hinton G.E.: Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process Syst.*, 2012. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf (dostęp: 25.11.2024)
- Simonyan K., Zisserman A.: Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv*, 2014; DOI: 10.48550/arXiv.1409.1556
- Szegedy Ch., Liu W., Jia Y. i wsp.: Going deeper with convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; DOI: 10.48550/arXiv.1409.4842
- He K., Zhang X., Ren S., Sun J.: Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf (dostęp: 25.11.2024)
- Krenn M., Buffoni L., Coutinho B. i wsp.: Predicting the future of ai with AI: high-quality link prediction in an exponentially growing knowledge network. *arXiv preprint arXiv*, 2022; DOI: 10.48550/arXiv.2210.00881
- Esteva A., Kuprel B., Novoa R.A. i wsp.: Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 2017; 542: 115–118
- De Fauw J., Ledsam J.R., Romera-Paredes B. i wsp.: Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat. Med.*, 2018; 24: 1342–1350
- Johnson A.E.W., Lehman L.W., Feng M. i wsp.: MIMIC-III, a freely accessible critical care database. *Sci. Data*, 2016; DOI: 10.1038/sdata.2016.35
- Johnson A.E.W., Pollard T.J., Berkowitz S.J. i wsp.: MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci. Data*, 2019; DOI: 10.1038/s41597-019-0322-0
- Irvin J., Rajpurkar P., Ko M. i wsp.: Chexpert: a large chest radiograph dataset with uncertainty labels and expert comparison. Proceedings of the AAAI Conference on Artificial Intelligence, 2019; 33: 590–597
- Grant D., Papięz B.W., Parsons G. i wsp.: Deep learning classification of cardiomegaly using combined imaging and non-imaging ICU data. Medical Image Understanding and Analysis: 25th Annual Conference, MIUA 2021, Oxford, United Kingdom, July 12–14, 2021, Proceedings 25. Springer International Publishing, 2021
- Hinton G.: On radiology. 24.11.2016. www.youtube.com/watch?v=2HMPRXtsVq (dostęp: 25.11.2024)
- Musk E.: Code conference. www.youtube.com/watch?v=wsixRS1-Sz4 (dostęp: 25.11.2024)
- Arterys 2017. www.forbes.com/sites/bernardmarr/2017/01/20/first-fda-approval-for-clinical-cloud-based-deep-learning-in-healthcare/ (dostęp: 25.11.2024)
- A million reasons for talking to Professor Martin Landray Landray 2021. University of Oxford. 20.05.2021. www.ox.ac.uk/news-and-events/oxford-people/martin-landray (dostęp: 25.11.2024)
- Roberts M., Driggs D., Thorpe M. i wsp.: Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat. Mach. Intell.*, 2021; 3: 199–217
- Data science and AI in the age of COVID-19 – report. 2020. www.turing.ac.uk/news/publications/data-science-and-ai-age-covid-19-report (dostęp: 25.11.2024)
- Ng A.Y., Glocker B., Oberije C. i wsp.: Artificial intelligence as supporting reader in breast screening: a novel workflow to preserve quality and reduce workload. *J. Breast Imag.*, 2023; 5: 267–276