

Uczenie maszynowe mądrzejsze, a nie tylko trudniejsze: Nowe ścieżki dla sztucznej inteligencji

Współczesna sztuczna inteligencja (ang. *Artificial Intelligence*, AI) potrafi dokonywać niesamowitych rzeczy, od rozpoznawania mowy, poprzez analizę tekstu i obrazów, po pomoc w diagnozowaniu chorób. Jednak proces „uczenia się” maszyn często bywa mało efektywny i nie przypomina sposobu, w jaki uczą się ludzie. Maszyny otrzymują przykłady uczące w losowej kolejności i na tej podstawie muszą budować ogólne rozumienie problemu. Niniejszy projekt ma na celu zmianę takiego podejścia poprzez opracowanie nowatorskich metod tzw. uczenia uporządkowanego (ang. *Curriculum Learning*, CL). Chcemy, aby komputery uczyły się podobnie jak dzieci, zaczynając od prostszych zadań i stopniowo przechodząc do coraz trudniejszych. Głównym celem projektu jest stworzenie niezawodnych, adaptacyjnych i uniwersalnych strategii uczenia maszynowego, które znacząco poprawią efektywność i wyniki działania modeli AI.

Małymi krokami ku mądrzejszej AI: Obecnie komputery często uczą się w sposób chaotyczny, próbując zrozumieć wszystkie dane jednocześnie, co może prowadzić do błędów lub „utknięcia” w martwym punkcie. Nasz projekt proponuje zintegrowane podejście, łączące trzy innowacyjne kierunki badań, aby proces uczenia maszyn był bardziej naturalny i skuteczny:

1. **Dynamiczne cele uczenia:** Wyobraźmy sobie, że uczymy maszynę rozpoznawać obrazy. Zamiast od razu pokazywać jej skomplikowane sceny, najpierw nauczymy ją rozróżniać ogólne kształty (np. czy to zwierzę, czy budynek), a dopiero potem przechodzimy do identyfikacji konkretnych detali (np. rasa psa, styl architektoniczny). W ramach projektu będziemy rozwijać metody, w których komputer na kolejnych etapach nauki skupia się na różnych aspektach problemu, stopniowo budując coraz pełniejsze zrozumienie. Wstępne wyniki, np. w zadaniu rozpoznawania dźwięków mowy, gdzie model najpierw uczył się ogólnych kategorii dźwięków (samogłoski, spółgłoski), a następnie poszczególnych fonemów, są bardzo obiecujące, pokazują szybsze dochodzenie do prawidłowych wyników i lepsze działanie na nowych danych.
2. **Uczenie kontekstowe i adaptacyjne:** Chcemy, aby maszyna, ucząc się nowej informacji, brała pod uwagę jej „sąsiedztwo”, czyli inne, podobne dane. Dodatkowo, sposób oceniania „błędów” popełnianych przez maszynę podczas nauki będzie dynamicznie dostosowywany. Na początku nauki system będzie mógł być traktowany łagodniej, zwłaszcza przy trudnych przykładach, a w miarę postępów „poprzeczka” uczenia będzie podnoszona.
3. **Inteligentne „mózgi” (sieci Kołmogorowa-Arnolda):** W projekcie wykorzystamy nowy typ sieci neuronowych, tzw. sieci Kołmogorowa-Arnolda. Ich unikalną cechą są „uczące się” funkcje aktywacji, czyli fundamentalne elementy przetwarzające informacje w sieci neuronowej, które mogą dynamicznie dostosowywać swoją złożoność. Planowane badania obejmą opracowanie metod „rozwój” tych funkcji tak, aby na początku model AI miał prostszą strukturę do nauki łatwiejszych zadań, a jego zdolności rosły wraz ze wzrostem trudności problemu.

Dlaczego podejmujemy ten temat badawczy? Tradycyjne metody uczenia uporządkowanego, mimo teoretycznych zalet, często zawodzą w praktyce, ponieważ są bardzo wrażliwe na ustawienia i rodzaj problemu. W dobie coraz większych i bardziej złożonych modeli AI (takich jak modele napędzające chatboty), potrzebujemy inteligentniejszych i bardziej wydajnych strategii ich trenowania. Obecne podejścia często wymagają ogromnych zasobów obliczeniowych i energii, co jest nie tylko kosztowne, ale też nieekologiczne. Opracowanie skuteczniejszych metod CL pozwoli na szybsze wdrażanie innowacji, zmniejszenie kosztów i zużycia energii oraz uczyni zaawansowaną AI bardziej dostępną dla mniejszych ośrodków badawczych i firm, które nie dysponują nieograniczonymi zasobami obliczeniowymi.

Jakich istotnych wyników oczekujemy? Realizacja projektu przyniesie szereg korzyści dla nauki oraz społeczeństwa. Oczekujemy, że opracowane metody pozwolą na trenowanie skuteczniejszych modeli AI, które będą osiągać lepsze wyniki, będą popełniać mniej błędów i będą lepiej radzić sobie z nowymi danymi. Nowe strategie uczenia uporządkowanego mają potencjał skrócenia czasu potrzebnego na wytrenowanie zaawansowanych modeli, co przełoży się na oszczędność zasobów obliczeniowych i energii. Modele trenowane z wykorzystaniem tych metod powinny być bardziej odporne na dane złej jakości, szum informacyjny czy próby celowego wprowadzenia modeli AI w błąd. Dodatkowo, projekt dostarczy naukowcom i inżynierom nowych, bardziej niezawodnych narzędzi do projektowania procesów uczenia maszynowego, co może przyspieszyć postęp w wielu dziedzinach nauki i technologii. Badania te przyczynią się również do głębszego zrozumienia dynamiki uczenia się złożonych modeli AI.