

Olimpiada Sztucznej Inteligencji

Witold DRZEWAŁOWSKI^{1,*,\dagger}, Kamil KSIĄŻEK^{2,\dagger},
Michał MARCINKOWSKI^{3,*}, Paulina TOMASZEWSKA^{4,*,\dagger}

¹Wydział Matematyki, Informatyki i Mechaniki, Uniwersytet Warszawski

²Wydział Matematyki i Informatyki, Uniwersytet Jagielloński

³Wydział Matematyki i Informatyki, Uniwersytet Wrocławski

⁴Wydział Matematyki i Nauk Informatycznych, Politechnika Warszawska

*Komitet Główny OAI

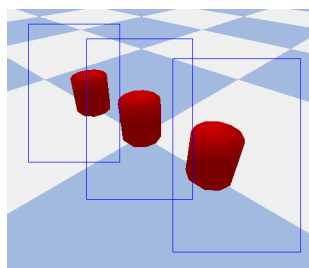
^{\dagger}Komitet Merytoryczny OAI

Sztuczna inteligencja coraz śmielej wkracza w nasze codzienne życie – modele językowe są podstawą działania chatbotów, silniki grające w szachy pomagają zarówno w szkoleniu początkujących graczy, jak i arcymistrzów, systemy wnioskowania rozwiązują skomplikowane zadania matematyczne. Za tym wszystkim stoją różnorodne algorytmy uczenia maszynowego i ogromne ilości danych. Jak to wszystko działa? Co stoi u podstaw tych systemów? Jak są one zbudowane? Olimpiada Sztucznej Inteligencji (OAI) została stworzona właśnie dla osób zainteresowanych tym obszarem wiedzy i technologii.

Co to jest OAI? Olimpiada Sztucznej Inteligencji to ogólnopolskie zawody adresowane do uczniów szkół ponadpodstawowych i podstawowych. Obecnie trwa III edycja Olimpiady, kolejna rozpocznie się w roku szkolnym 2026/2027. Wzorem innych olimpiad, takich jak matematyczna i informatyczna, OAI składa się z trzech etapów: I etapu zdalnego, II etapu lokalnego oraz III etapu finałowego. W roku szkolnym 2025/2026 I etap rozpoczął się 1 grudnia 2025 roku i zakończy się 25 stycznia 2026, II etap odbędzie się w dniach 6–8 marca 2026 roku w czterech miastach: Krakowie, Poznaniu, Warszawie i Wrocławiu. Finał będzie miał miejsce od 17 do 19 kwietnia 2026 roku w Poznaniu. OAI jest wspierana przez Ministerstwo Edukacji Narodowej. Jej organizatorem jest Fundacja Edukacji i Rozwoju Sztucznej Inteligencji oraz wydziały matematyki i informatyki uniwersytetów Jagiellońskiego, im. Adama Mickiewicza, Warszawskiego i Wrocławskiego.

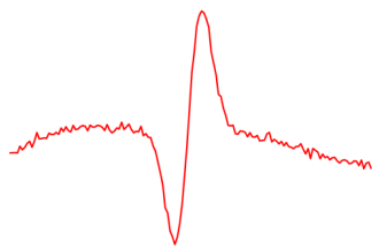
Jak wyglądają zadania? Zadania inspirowane są realnymi problemami pojawiającymi się w przemyśle i nauce. Uczestnicy muszą przeanalizować duży zbiór danych i znaleźć w nim ukryte zależności, tworząc własny model sztucznej inteligencji, np. sieć neuronową. W poprzednich edycjach zadania dotyczyły m.in. wykrywania nieprawidłowości w sygnale EKG (ilustrującym pracę serca), detekcji halucynacji w danych wygenerowanych przez duże modele językowe, czy predykcji sumy pieniędzy na zdjęciu z monetami. Jedno z zadań związane było z analizą wideo przedstawiającą grę w trzy kubki. Na filmach kubki były zamieniane miejscami, a zadaniem było wskazanie, na jakiej pozycji finalnie znalazł się każdy kubek. Kolejne zadanie dotyczyło ataków na modele AI, aby „zmusić” je do popełniania błędów. Aby lepiej zilustrować, na czym mogą polegać zadania, opiszemy bardziej szczegółowo dwa z nich.

- **Wykrywanie zaburzeń sygnału EKG** (II OAI, 1. etap). W tym zadaniu uczestnik dostaje zbiór danych zawierający syntetyczne sygnały EKG, czyli zapis pracy serca w czasie. Część sygnałów EKG jest prawidłowa, a część odpowiada wybranym schorzeniom, np. migotaniu przedsionków. Zadaniem uczestnika jest wyodrębnienie z fragmentów EKG czterech liczbowych cech sygnału, które pozwolą najlepiej przewidzieć obecność schorzeń. Dobór cech jest oceniany przez jakość wytrenowanego na ich podstawie modelu uczenia maszynowego, którego uczestnicy nie mogą zmieniać. Im lepsze cechy, tym dokładniejsze przewidywania modelu oraz wyższa ocena rozwiązania. Zadaniem uczestników było odkrycie anomalii, które znajdowały się tylko w szczególnych fragmentach sygnału. Jedno z rozwiązań polegało na podzieleniu sygnału EKG na fragmenty i usunięciu z niego odcinków odpowiadających tzw. zespołowi QRS, reprezentującemu główną aktywność skurczową serca. Pozostałe fragmenty, odpowiadające sygnałowi pomiędzy kolejnymi uderzeniami serca, zawierały większość anomalii i były podstawą dalszej analizy statystycznej. Badano między innymi wartości bezwzględne różnic między kolejnymi odczytami w czasie, a ostateczne cechy były konstruowane na podstawie ich odchylenia standardowego, średniej kroczącej, amplitudy lub innych statystyk. Zaznaczamy, że rozwiązanie nie wymagało wiedzy medycznej, a całe niezbędne wprowadzenie, włącznie z opisem zespołu QRS, było przedstawione w treści zadania.

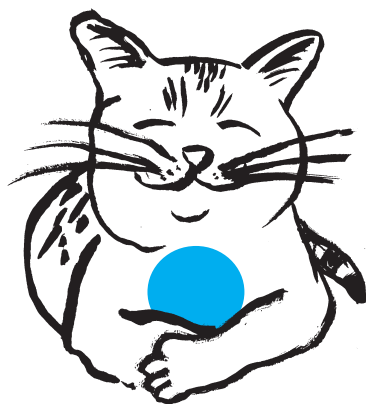


Zadanie z grą w trzy kubki wymagało śledzenia obiektów występujących w filmie

Halucynacje AI to zjawisko, w którym sztuczna inteligencja generuje fałszywe lub nieprawdziwe informacje, prezentując je jako fakty. Powstają one, ponieważ modele AI przewidują najbardziej prawdopodobne odpowiedzi na podstawie wzorców w danych treningowych, a nie na prawdziwym zrozumieniu rzeczywistości, co może prowadzić do zmyślenia faktów, dat, a nawet źródeł.



Sygnał EKG ze zbioru treningowego



W powyższym zadaniu nacisk postawiony był na analizę danych wejściowych. Model nie był modyfikowany przez użytkownika. Nie zawsze jednak tak jest. Poniżej opiszemy zadanie z finałowego etapu, w którym należało dotrenować dostarczony model sieci neuronowej.

- **Stylizacja tłumaczeń maszynowych** (II OAI, 3. etap). Uczestnicy w ramach zadania dostali wytrenowany już model niewielkich rozmiarów do tłumaczeń z języka angielskiego na język polski oraz zdania w języku angielskim zawierające specjalistyczne terminy z dziedziny uczenia maszynowego. Dla każdego z takich zdań zbiór danych zawierał również listę występujących w nich specjalistycznych terminów oraz oczekiwane tłumaczenie tego zdania na język polski. Jednakże w tym tłumaczeniu terminy specjalistyczne są przetłumaczone inaczej niż dokonałby tego dostarczony model – stosowane są polskie odpowiedniki, podczas gdy celem tego zadania było pozostawienie angielskich terminów. Zadanie to wymagało implementacji pętli do wydajnego dotrenowania modelu z uwzględnieniem nowoczesnych technik, takich jak zmiana tempa treningu w ramach kolejnych etapów poprzez zmniejszanie współczynnika uczenia, z uwzględnieniem fazy „rozgrzewki”. Takie rozwiązanie często nie skutkowało jednak maksymalną liczbą punktów. Kluczowa była następująca obserwacja: aby skutecznie naśladować styl tłumaczeń obecny w zbiorze danych, model musi opanować dwie umiejętności: (1) rozpoznawanie fraz specjalistycznych oraz (2) odpowiednie tłumaczenie tych fraz. Uczestnicy, którzy to dostrzegli, wykorzystali dodatkową informację dostępną w danych – listy słów kluczowych przypisane do każdego zdania. Na ich podstawie oznaczali w tekście źródłowym (angielskim) frazy specjalistyczne (np. za pomocą znaczników `<...>`). Dzięki temu model nie musiał samodzielnie uczyć się, które wyrażenia wymagają zmiany stylu, a które nie. Jego zadaniem było jedynie nauczyć się, że frazy oznaczone należy tłumaczyć zgodnie ze specyficznym stylem, a w tłumaczeniu pomijać same znaczniki.

Wiele zadań wymaga obliczeń na kartach graficznych (GPU), do których zapewniamy dostęp podczas etapów stacjonarnych. Zadania są w pełni automatycznie oceniane na specjalnie przygotowanej infrastrukturze systemu sprawdzającego.

Dlaczego warto? Poza satysfakcją i świetną przygodą Olimpiada to szansa na rozwinięcie bardzo przydatnych umiejętności oraz możliwość spotkania rówieśników z całej Polski zainteresowanych sztuczną inteligencją. Na etapy lokalne i finał zapraszamy naszych partnerów i firmy wspierające Olimpiadę – jest wtedy okazja do uczestniczenia w ciekawych wykładach i zobaczenia, jak AI jest stosowana w praktyce. Finaliści mają ułatwiony lub wolny wstęp na coraz większą liczbę kierunków na polskich uczelniach. Dla najlepszych uczestników organizujemy obóz naukowy, którego poprzednie edycje odbywały się w Krzyżowej i Polanicy-Zdroju. Uczestnicy poznają tam bardziej zaawansowane zagadnienia, takie jak modele dyfuzyjne w wizji komputerowej, sieci typu transformer czy metody trenowania modeli językowych. Co roku Olimpiada wyłania polską reprezentację na zawody międzynarodowe. Dotychczas odbyły się dwie edycje Międzynarodowej Olimpiady Sztucznej Inteligencji (ioai-official.org). W obu tych olimpiadach nasza reprezentacja osiągnęła wspaniałe sukcesy. Na ostatniej międzynarodowej olimpiadzie w Pekinie polska reprezentacja zdobyła 7 medali (3 złote, 3 srebrne, 1 brązowy), zajmując w nieoficjalnej klasyfikacji medalowej **2 miejsce na świecie**.

Jak zacząć? Wszystkie potrzebne informacje, materiały przygotowujące do Olimpiady, a także zadania z poprzednich edycji znajdują się na stronie oai.edu.pl oraz w mediach społecznościowych Olimpiady. W ramach Olimpiady prowadzimy również różnorodne szkolenia. Przed I etapem organizujemy cykl wykładów wprowadzających, podczas których od podstaw przedstawiamy zagadnienia z zakresu uczenia maszynowego. Omawiamy m.in. podstawy uczenia maszynowego, elementy wizji komputerowej i przetwarzania języka naturalnego. Wprowadzamy uczestników do programowania w Pythonie, trenujemy pierwsze sieci neuronowe, pokazujemy algorytmy klasteryzacji, sieci konwolucyjne do rozpoznawania obrazów oraz wektory zanurzeń słów Word2Vec.

Warto podkreślić, że dla większości naszych uczestników start w Olimpiadzie był ich pierwszą przygodą z AI – jesteśmy jednak przekonani, że jest to dopiero początek ich fascynującej podróży.

