

Streszczenie

W niniejszej rozprawie doktorskiej przedstawiono cykl trzech publikacji naukowych poświęconych tematyce uczenia ciągłego (ang. Continual Learning) w systemach sztucznej inteligencji. Głównym wyzwaniem badawczym podjętym w pracy jest realizacja procesu uczenia w warunkach całkowitego braku dostępu do danych historycznych. Oznacza to, że podczas adaptacji modelu do nowego zadania zabronione jest korzystanie z jakichkolwiek próbek danych dostępnych na wcześniejszych etapach szkolenia. Jest to najbardziej rygorystyczny i wymagający wariant uczenia przyrostowego, w którym sieci neuronowe wykazują naturalną tendencję do tzw. katastroficznego zapominania (ang. catastrophic forgetting). Zjawisko to polega na gwałtownej utracie zdolności do rozwiązywania wcześniej opanowanych zadań na skutek nadpisywania wag modelu nowymi informacjami. Zaprezentowane publikacje analizują i generalizują problem uczenia ciągłego, a także proponują nowatorskie metody jego rozwiązania.

W pierwszej publikacji przedstawiono nową metodę uczenia zespołowego, w której eksperci współpracują ze sobą przy rozwiązywaniu zadań i autonomicznie decydują, który z nich zostanie poddany procesowi douczania do nowego zadania. Dzięki zastosowaniu heurystyki analizującej rozkłady wiedzy ekspertów, do adaptacji wybierany jest ten model, którego stopień zapominania będzie najmniejszy. W kolejnej publikacji problematyka uczenia ciągłego została rozszerzona o scenariusze, w których zbiory danych są tylko częściowo etykietowane. Wykazano, że połączenie destylacji wiedzy z adaptacją klas jest kluczowe dla uzyskania wysokiej plastyczności modelu przy jednoczesnym ograniczeniu zapominania. W ostatniej, trzeciej publikacji pokazano, że powszechny problem faworyzacji ostatnich zadań (ang. task recency bias) wynika z zapaści wymiarowej (ang. dimensional collapse) w przestrzeni reprezentacji (zanurzeń) tych modeli. W odpowiedzi na to zjawisko opracowano nowatorską funkcję straty wymuszającą dodatnie wartości własne macierzy kowariancji klas oraz algorytm ich adaptacji, co znacząco podnosi efektywność procesów uczenia ciągłego w złożonych architekturach głębokich.

Słowa kluczowe: Sztuczna Inteligencja, Uczenie głębokie, Uczenie Ciągłe, Katastroficzne Zapominanie

Publikacje wchodzące w skład rozprawy doktorskiej

#	Tytuł	Autorzy	Konferencja	Status
1.	Divide and not forget: Ensemble of selectively trained experts in Continual Learning	G. Rypeś, S. Cygert, V. Khan, T. Trzciński, B. Zieliński, B. Twardowski	ICLR 2024	Zaakceptowana
2.	Category Adaptation Meets Projected Distillation in Generalized Continual Category Discovery	G. Rypeś, D. Marczak, S. Cygert, T. Trzciński, B. Twardowski	ECCV 2024	Zaakceptowana
3.	Task-recency bias strikes back: Adapting covariances in Exemplar-Free Class Incremental Learning	G. Rypeś, S. Cygert, T. Trzciński, B. Twardowski	NeurIPS 2024	Zaakceptowana