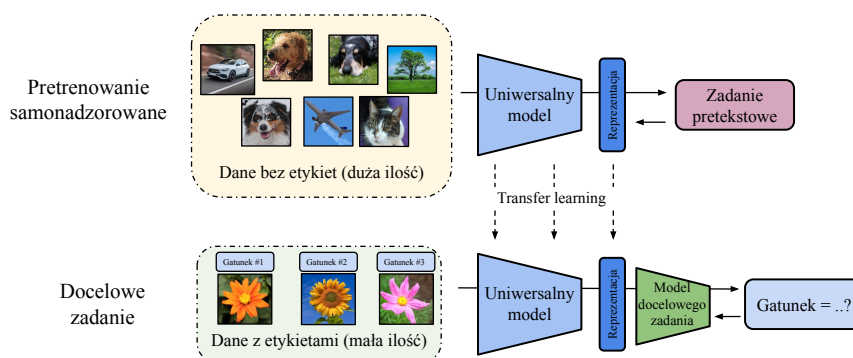


Usprawnienie adaptowalności modeli uczonych metodami samonadzorowanymi

Nr rejestracyjny: 2023/49/N/ST6/03268; Kierownik projektu: mgr inż. Marcin Paweł Przewiński

W ciągu ostatniej dekady ludzkość była świadkiem bezprecedensowego wzrostu możliwości modeli głębokiego uczenia. W rezultacie, sieci neuronowe zostały powszechnie przyjęte jako najnowocześniejsze podejście do sztucznej inteligencji i obecnie codziennie usprawniają pracę, rozrywkę i opiekę zdrowotną miliardów ludzi. Jednak głębokie sieci neuronowe polegają na dużych ilościach danych, aby móc nauczyć się wykonywać swoje zadania. Stanowi to wąskie gardło we wdrażaniu modeli głębokiego uczenia w rzeczywistych problemach — oznaczenie wystarczającej ilości danych, aby wyszkolić model do nowego zadania od podstaw w czysto nadzorowany sposób, jest często trudne lub kosztowne. Na szczęście uczenie nadzorowane nie jest jedynym sposobem uczenia głębokich sieci neuronowych.

Uczenie samonadzorowane (ang. *Self-supervised learning*) to paradygmat uczenia się reprezentacji danych bez etykiet, z zamiarem późniejszego *przeniesienia* tych reprezentacji (ang. *transfer learning*) do określonych zadań określonych przez dane oetykietowane (Rysunek 1). W praktyce, dostrojenie takich pretrenowanych modeli wymaga znacznie mniejszych zestawów danych z etykietami oraz zasobów obliczeniowych w porównaniu do uczenia ich od podstaw. Uczenie samonadzorowane otwiera możliwości pretrenowania modeli na ogromnych ilościach nieoznakowanych danych, co pozwoliło na sukcesy w różnych dziedzinach, od przetwarzania języka naturalnego, aż po wizję komputerową.



Rysunek 1: Uczenie samonadzorowane trenuje głębokie sieci neuronowe do tworzenia *ekspresywnych reprezentacji danych* poprzez rozwiązywanie *zadań pretekstowych*, które nie wymagają etykietowania danych uczących. Następnie pretrenowany model może być ponownie wykorzystany jako fundament dla modeli rozwiązujących różne rodzaje docelowych zadań opisanych przez dane z etykietami. Ponieważ przenosimy wstępnie wytrenowany model, nie potrzebujemy dużej ilości danych z etykietami, aby dostroić go do określonych zadań.

Reprezentacje trenowane przez uczenie samonadzorowane nie są jednak wolne od typowych problemów, które mogą wystąpić podczas uczenia z przenoszeniem reprezentacji. Przykładowo, przenoszony model dziedziczy nie tylko wytrenowaną reprezentację, ale także uprzedzenia nauczone podczas nauki zadania użytego podczas pretrenowania. Na przykład kontrastowy cel niezmienniczości na zmiany w kolorach (ang. *color jittering*) spowoduje, że model będzie radzić sobie źle w zadaniach zależnych od tej cechy danych, takich jak rozpoznawanie gatunków kwiatów.

Powyższe możliwości i ograniczenia uczenia samonadzorowanego inspirują nas do skupienia się w tym projekcie badawczym na **usprawnieniu adaptowalności modeli uczonych metodami samonadzorowanymi**. Głównym kierunkiem badawczym, który chcielibyśmy obrać, jest problem łagodzenia skutków uprzedzeń indukowanych w reprezentacjach przez określone funkcje kosztu stosowane w czasie pretrenowania, takich jak niezmienniczość na augmentacje. Oczekiwane wyniki naszych badań mogą potencjalnie wpłynąć na inne dziedziny uczenia maszynowego, w których adaptowalność jest istotną kwestią, takie jak Meta-Uczenie (ang. *Meta-learning*) oraz Uczenie Ciągłe (ang. *Continual Learning*).

W naszych eksperymentach wykorzystamy szeroką gamę współczesnych technik głębokiego uczenia, takich jak warunkowanie sieci neuronowych różnymi modalnościami danych, hipersieci, transformery i mechanizmy uwagi. Przyjrzymy się koncepcji warunkowania sieciowego inspirowanej mechanizmem uczenia generatywnych modeli warunkowych, który nie był wcześniej stosowany w ramach paradygmatu uczenia samonadzorowanego.

Mamy nadzieję, że niniejszy projekt badawczy zaowocuje odkryciem nowych faktów na temat uczenia samonadzorowanego i udoskonalonych sposobów uczenia bardziej uniwersalnych modeli podstawowych, które można w prosty sposób wykorzystać do wielu zadań, umożliwiając zastosowanie głębokiego uczenia w większej liczbie rzeczywistych problemów.