

Głębokie nauczanie w reprezentacji długich danych sekwencyjnych

Popularnonaukowy opis badań

Konrad Żoła, grudzień 2017

Umiejętność wydobycia z pamięci ważnych informacji i doświadczeń pomimo upływu czasu pozwala nam podjąć trafniejsze decyzje. Innymi słowy, rzetelne przedstawienie obecnej sytuacji wymaga wzięcia pod uwagę zgromadzonych wcześniej informacji oraz poprzednio wykonanych akcji, mających wpływ na obecny stan. Odchodząc od tak ekstremalnych przykładów, nawet pozornie niewymagające zadanie streszczenia filmu wymaga od nas odpowiedniego filtrowania treści i zapamiętywania najważniejszych informacji podczas całego procesu oglądania. Człowiek wspaniale radzi sobie z tym zadaniem i nawet nieuprzedzony będzie w stanie opisać film bez większych kłopotów.

Aby maszyny były zdolne do rozwiązywania takich zadań konieczne jest stworzenie architektury, która potrafi opisać swój stan w obecnej sytuacji biorąc pod uwagę wszystkie najważniejsze informacje, nawet te uzyskane bardzo dawno. Niestety istniejące obecnie modele mają spore trudności ze znajdowaniem rozległych w czasie związków przyczynowo-skutkowych. Dlatego też, celem moich badań jest przygotowanie nowej architektury sztucznej sieci neuronowej, która będzie potrafiła opisać i streścić długą sekwencję historycznych obserwacji. A zatem podobnie jak człowiek, będzie skupiać się na najważniejszych bodźcach i pamiętać je tak długo, jak długo pozostają w grupie najistotniejszych czynników kształtujących aktualną sytuację.

Pojęcie sztucznych sieci neuronowych stanowi bardzo ogólne określenie obiektów matematycznych, których oryginalną inspiracją była budowa mózgu. Praktyka pokazała, że ta metoda jest bardzo uniwersalna i może zostać z powodzeniem użyta do rozwiązania wielu rodzajów problemów, jak na przykład rozpoznawanie obrazów, synteza mowy, czy wykrywanie chorób nowotworowych. Sztuczne sieci neuronowe są ostatnio bardzo intensywnie badane. Opisano specjalne wersje, które potrafią analizować sekwencje obserwacji i dlatego wierzę, że ta metoda jest najbardziej obiecującym rozwiązaniem mojego problemu badawczego.

Powodem podjęcia danej tematyki badawczej jest nie tylko olbrzymie zainteresowanie tym tematem w świecie uczonych zajmujących się zagadnieniem sztucznej inteligencji, ale także szereg potencjalnych zastosowań. Oto tylko kilka przykładów obrazujących praktyczne problemy, które potencjalnie wykorzystują badany przeze mnie problem.

- Reprezentacja dokumentów tekstowych pozwoliłaby przygotować odpowiedzi na zadane przez człowieka pytania w języku naturalnym (źródłem wiedzy byłyby korpusy spisane także w języku naturalnym, jak na przykład encyklopedie), bądź przygotowanie automatycznych streszczeń bardzo długich dokumentów.
- Wyznaczenie stanu przemieszczającej się maszyny lub kamery, w celu skonstruowania mapy okolicznego terenu. Taka mapa może zostać wykorzystana na przykład do generowania instrukcji dla niewidomych.
- Reprezentacja użytkownika strony internetowej, w celu przygotowania lepszych rekomendacji. Bezpośrednim skutkiem byłoby tu przyspieszenie składania zamówień i w efekcie skrócenie tak powtarzalnej czynności, jaką są zakupy.

Praca na tym problemem jest szczególnie ciekawa także dlatego, że nawet częściowe rozwiązania znajdują zastosowanie. Jednym z przykładów jest metoda (powstała jako rezultat badań wstępnych), która pozwala zidentyfikować, które z badanych dzieci w wieku szkolnym mają problemy z pisaniem i jeżeli interwencja jest potrzebna, szybciej zaadresować ten problem.