

Wielokryterialna modularność w przetwarzaniu języka naturalnego

Dziedzina przetwarzania języka naturalnego (NLP) odnotowała w ostatnich latach błyskawiczny postęp dzięki ciągłemu skalowaniu głębokich sieci neuronowych (DNN). W rezultacie, pojawienie się modeli podstawowych znacząco zmieniło krajobraz NLP, gdzie wstępnie wytrenowane modele językowe i duże modele językowe (LLM) osiągają wyjątkowe rezultaty w licznych zadaniach.

Trwający proces skalowania głębokich sieci neuronowych doprowadził do wzrostu popularności modularnego uczenia głębokiego (MDL). Metody MDL prezentują tańszą alternatywę do całkowitego dostrajania modeli poprzez parametrycznie efektywny odpowiednik, który redukuje wymagania obliczeniowe modyfikacji modelu. W szczególności, w przypadku sztandarowych modeli LLM, które swoim rozmiarem przekraczają 100 miliardów parametrów, nawet predykcja wymaga znacznej infrastruktury sprzętowej, pomijając całkowicie koszty treningu takiego modelu. Co więcej, moduły MDL mogą wzbogacać modele podstawowe dodając im różne umiejętności, często wykorzystując jednocześnie wiele z nich i łącząc je w zależności od potrzeb zadania docelowego.

Niniejsza praca skupia się na modularnym uczeniu głębokim i bada wielokryterialne aspekty metod modularnych w formie serii czterech publikacji. W pierwszych trzech z nich stawiane jest pytanie badawcze, które rozważa konkretny problem w istniejących aspektach: wielozadaniowości, wielodomenowości i wielojęzyczności. Owe prace skupiają się na jednym, konkretnym przypadku wielokryteriowości na raz, bez modyfikacji pozostałych (np. podczas eksploracji konfiguracji zadań wielozadaniowych zachowujemy niezmienną domenę i język(i)). Każda z pierwszych trzech publikacji przedstawia nowy system lub metodę, która ma na celu usunięcie ograniczeń obecnych rozwiązań w danym aspekcie. W ostatniej, czwartej publikacji, został sformułowany nowy aspekt, nazwany wielomodelowym. Praca testuje istniejące metody modularne w nowych warunkach i ukazuje ograniczenia obecnych metod modularnych.

Słowa kluczowe: Modularne uczenie głębokie, Parametrycznie efektywne dostrajanie, Przetwarzania języka naturalnego