

Omówienie aktualnych postępów w dziedzinie ciągłego uczenia maszynowego

Monika Etrych

15 czerwca 2024

1 Wprowadzenie

Ciągłe uczenie maszynowe (continuous learning, lifelong learning, incremental learning, sequential learning) to metoda trenowania modeli sztucznej inteligencji, w której model jest regularnie aktualizowany na podstawie nowych danych napływających w czasie rzeczywistym. Umożliwia to dynamiczne dostosowywanie się modelu do zmieniających się warunków bez konieczności przeprowadzania pełnego procesu trenowania od podstaw.

Metody uczenia ciągłego można podzielić na uczenie online i offline. Uczenie online polega na bieżącej aktualizacji modelu po otrzymaniu każdej nowej próbki danych, co pozwala na natychmiastową adaptację do nowych informacji. Jest to szczególnie użyteczne w dynamicznych środowiskach, gdzie dane są stale strumieniowane. Z kolei uczenie offline, nazywane także batch learning, obejmuje okresowe aktualizacje modelu na podstawie zebranych danych. Dane są gromadzone przez pewien czas, a następnie wykorzystywane do retreningu modelu. Ta metoda jest wykorzystywana w sytuacjach, gdzie dostępność zasobów obliczeniowych jest ograniczona lub gdy aktualizacje nie mogą być dokonywane tak często.

2 Wyzwania i metody ich rozwiązywania

Kluczowe wyzwania ciągłego uczenia maszynowego obejmują efektywne zarządzanie zasobami obliczeniowymi, zapobieganie zapominaniu wcześniej nabytej wiedzy (catastrophic forgetting) oraz zapewnienie stabilności i plastyczności modelu (stability-plasticity dilemma), czyli jednoczesnego pamiętania zdobytej wiedzy oraz nauki nowych danych.

W ciągłym uczeniu maszynowym występują interferencje między nowymi a wcześniejszymi zadaniami, które często prowadzą do katastrofalnego zapominania, gdzie model traci zdolność do przypominania sobie wcześniej nabytych informacji. Jest to wynik zmian wag modelu, dokonywanych podczas uczenia się nowych danych, które wpływają na wcześniej nauczone informacje.

Aby minimalizować problem zapominania, stosuje się różne metody, takie jak regularyzacja, izolacja parametrów, modyfikacje architektury oraz metody oparte na powtarzaniu danych. Metody regularyzacyjne mają na celu penalizowanie dużych zmian wag, co stabilizuje model i zapobiega zapominaniu wcześniej nauczonych danych. Metody izolacji parametrów, takie jak Synapsis Intelligence, śledzą istotność poszczególnych wag i

ograniczają ich zmiany, aby chronić istotne informacje. Inne podejścia, jak Progress nad Compress, dzielą parametry odpowiedzialne za różne zadania, aby chronić ich specyficzność. Dostosowanie architektury poprzez dodanie nowych neuronów lub warstw również może zapobiec nadpisaniu istotnych wag. Metody oparte na powtarzaniu danych pozwalają na przypomnienie wcześniejszych zadań, aby wagi odpowiedzialne za te zadania nie zostały zapomniane, co może być osiągnięte poprzez przechowywanie danych z wcześniejszych zadań lub generowanie ich na nowo.

3 Przegląd literatury

Artykuł „*Continual Learning with Knowledge Distillation: A Survey*” analizuje metody ciągłego uczenia maszynowego oparte na distylacji wiedzy, które redukują problem katastroficznego zapominania. Autorzy eksperymentalnie potwierdzają skuteczność distylacji wiedzy w kontekście klasyfikacji obrazów, takich jak CIFAR-100, Tiny-ImageNet i ImageNet-100, gdzie bardziej złożony model (student) jest uczony na podstawie wyjść mniej złożonego modelu (teacher). Distylacja to technika w której bardziej złożony model (student model) jest uczony na podstawie wyjść mniej złożonego modelu (teacher model).

Przegląd „*A Comprehensive Survey of Continual Learning: Theory, Method and Application*” przedstawia połączenie uczenia ciągłego z innymi zagadnieniami sztucznej inteligencji takimi jak: modele dyfuzyjne, modele fundamentowe, architektury oparte na transformerach, uczenie wielomodalne, sztuczna inteligencja osadzona.

W artykule „*Continual Lifelong Learning with Neural Networks: A Review*” dokonano porównania wyników wybranych modeli na różnych zbiorach danych obrazowych takich jak MNIST, CUB-200 i CORe50. Przedstawione metody to EWC (Elastic Weight Consolidation), FEL (Forget-Extra Learning), MLP (Multi-Layer Perceptron), GeppNet (Generalized Episodic Parameter Propagation Network) oraz GeppNet + STM (Short-Term Memory). Artykuł wskazuje, że GeppNet osiąga najlepsze wyniki na zbiorze MNIST, a GeppNet + STM na zbiorze CUB-200 oraz AudioSet.

4 Podsumowanie

Ciągłe uczenie maszynowe jest kluczowym obszarem badań w sztucznej inteligencji, umożliwiającym modelom dynamiczne dostosowywanie się do zmieniających się warunków i napływających danych. Dzięki metodom takim jak distylacja wiedzy, regularyzacja, izolacja parametrów, modyfikacje architektury oraz powtarzanie danych, można skutecznie minimalizować problem katastroficznego zapominania i zapewniać równowagę między stabilnością a plastycznością modeli. Przeglądy literatury wskazują na rosnące zainteresowanie ciągłym uczeniem w różnych dziedzinach AI, co otwiera nowe możliwości i wyzwania dla dalszych badań i aplikacji. W miarę jak technologia ta będzie się rozwijać, możemy spodziewać się coraz bardziej zaawansowanych i wszechstronnych systemów sztucznej inteligencji, zdolnych do efektywnego uczenia się przez całe życie.

5 Literatura

[1] - Continual Learning with Knowledge Distillation: A Survey, 2024, Songze Li, Tonghua Su Member, Xuyao Zhang Senior Member, Zhongjie Wang Member

[2] - A Comprehensive Survey of Continual Learning: Theory, Method and Application, 2024, Liyuan Wang, Xingxing Zhang, Hang Su, Jun Zhu, Fellow

[3] - Continual Lifelong Learning with Neural Networks: A Review, 2019, German I. Parisi¹, Ronald Kemker, Jose L. Part, Christopher Kanan, Stefan Wermter Knowledge Technology