

Ontologia wartości atrybutu jako przykład zastosowania semantyki w uczeniu maszynowym

dr inż. Tomasz Łukaszewski, prof. dr hab. inż. Joanna Józefowska, dr inż. Agnieszka Ławrynowicz

Abstrakt: Uczenie maszynowe jest dziedziną, w której osiągnięto już wiele istotnych wyników badawczych. W ostatnich kilku latach daje się jednak zauważyć prace badawcze, które T. Łukaszewski, J. Józefowska, A. Ławrynowicz Ontologia wartości atrybutu jako przykład zastosowania semantyki w uczeniu maszynowym.

Uczenie maszynowe jest dziedziną, w której osiągnięto już wiele istotnych wyników badawczych. W ostatnich kilku latach daje się jednak zauważyć prace badawcze, które podchodzą do zagadnienia uczenia maszynowego w sposób niestandardowy. Podejścia takie pozwalają rozwiązać pewne istotne problemy uczenia maszynowego, nie mające efektywnych rozwiązań w klasycznych metodach.

Przykładem takiego niestandardowego podejścia jest zastosowanie semantyki w uczeniu maszynowym, prezentowane m.in. w niniejszej pracy. Rozszerzając dziedzinę wartości atrybutów do ontologii wartości atrybutów, dodajemy do istniejących już wartości precyzyjnych (reprezentowane przez pojęcia prymitywne ontologii), wartości nieprecyzyjne (reprezentowane przez pojęcia abstrakcyjne ontologii). Wartości nieprecyzyjne mogą być na przykład zbiorami wartości precyzyjnych, w szczególności przedziałami liczbowymi. Ontologia wartości atrybutu wzbogaca język reprezentacji wiedzy i poprawne jej wykorzystanie pozwala zmierzyć się z problemami klasycznego uczenia maszynowego: reprezentacją nieprecyzyjności opisu, zbyt mocnym dopasowaniem modelu do zbioru uczącego (ang. overfitting) czy wysokimi kosztami uzyskania wartości precyzyjnych.

Przeprowadzone przez nas wstępne eksperymenty obliczeniowe wskazują, że trafność klasyfikacji nie ulegnie znacznemu pogorszeniu, gdy pozostawi się pewne atrybuty nieokreślone (tzw. wartość brakująca), a zaproponowany prototyp ontologicznego klasyfikatora bayesowskiego potwierdza możliwość poprawnego wykorzystania semantyki w uczeniu maszynowym. podchodzą do zagadnienia uczenia maszynowego w sposób niestandardowy. Podejścia takie pozwalają rozwiązać pewne istotne problemy uczenia maszynowego, nie mające efektywnych rozwiązań w klasycznych metodach.

Przykładem takiego niestandardowego podejścia jest zastosowanie semantyki w uczeniu maszynowym, prezentowane m.in. w niniejszej pracy. Rozszerzając dziedzinę wartości atrybutów do ontologii wartości atrybutów, dodajemy do istniejących już wartości precyzyjnych (reprezentowane przez pojęcia prymitywne ontologii), wartości nieprecyzyjne (reprezentowane przez pojęcia abstrakcyjne ontologii). Wartości nieprecyzyjne mogą być na przykład zbiorami wartości precyzyjnych, w szczególności przedziałami liczbowymi. Ontologia wartości atrybutu wzbogaca język reprezentacji wiedzy i poprawne jej wykorzystanie pozwala zmierzyć się z problemami klasycznego uczenia maszynowego: reprezentacją nieprecyzyjności opisu, zbyt mocnym dopasowaniem modelu do zbioru uczącego (ang. overfitting) czy wysokimi kosztami uzyskania wartości precyzyjnych.

Przeprowadzone przez nas wstępne eksperymenty obliczeniowe wskazują, że trafność klasyfikacji nie ulegnie znacznemu pogorszeniu, gdy pozostawi się pewne atrybuty nieokreślone (tzw. wartość brakująca), a zaproponowany prototyp ontologicznego klasyfikatora bayesowskiego potwierdza możliwość poprawnego wykorzystania semantyki w uczeniu maszynowym.