

Adam Szczur

Wybrane metody konstrukcji klasyfikatorów dla danych temporalnych

Streszczenie

Rozprawa poświęcona jest problematyce konstrukcji klasyfikatorów dla danych temporalnych reprezentowanych w postaci temporalnych tablic decyzyjnych. Dane tego typu występują powszechnie w wielu dziedzinach, takich jak medycyna, przemysł, sport czy systemy monitorowania procesów dynamicznych. W klasycznych podejściach do aproksymacji pojęć czasowych dane temporalne są najczęściej transformowane do postaci statycznych reprezentacji cech opisujących całe okna czasowe. Umożliwia to zastosowanie standardowych metod klasyfikacji, jednak prowadzi do uproszczenia informacji zawartej w przebiegu zmian obserwowanych w czasie. Alternatywę stanowią metody głębokiego uczenia, które umożliwiają bezpośrednie przetwarzanie danych temporalnych, jednak kosztem ograniczonej interpretowalności modeli.

Celem rozprawy było opracowanie interpretowalnych metod klasyfikacji danych temporalnych umożliwiających bezpośrednie wykorzystanie informacji zawartej w oknach czasowych bez konieczności ich wcześniejszej agregacji. W pracy zaproponowano autorską metodę TC-Tree (Decision Tree with Temporal Cuts), opartą na koncepcji temporalnych cięć, umożliwiających identyfikację zależności występujących wewnątrz okien czasowych i wykorzystanie ich do budowy modeli klasyfikacyjnych w postaci drzew decyzyjnych. Następnie przedstawiono metodę VTC-Tree (Decision Tree with Verifying Temporal Cuts), stanowiącą rozszerzenie podejścia TC-Tree poprzez wprowadzenie mechanizmu weryfikujących cięć temporalnych wspomagających proces podejmowania decyzji klasyfikacyjnych.

Dodatkowo w rozprawie przedstawiono metodę identyfikacji i przewidywania zmian wzorców zachowania opracowaną dla danych medycznych związanych z niewydolnością oddechową wcześniaków. Metoda ta stanowi odrębny komponent badawczy dotyczący analizy zmian temporalnych w danych rzeczywistych.

Ocena eksperymentalna zaproponowanych metod została przeprowadzona z wykorzystaniem danych pochodzących z różnych dziedzin zastosowań, obejmujących między innymi aktywność robotów, analizę zachowań sportowców oraz dane medyczne. Uzyskane wyniki wykazały, że metody TC-Tree i VTC-Tree osiągają skuteczność klasyfikacji porównywalną, a w wybranych przypadkach wyższą niż klasyczne metody klasyfikacji oraz wybrane metody głębokiego uczenia, takie jak sieci 1D-CNN i LSTM. Jednocześnie proponowane podejścia zachowują wysoki poziom interpretowalności modeli, umożliwiając analizę procesu podejmowania decyzji oraz identyfikację reguł odpowiedzialnych za klasyfikację obiektów.

Przeprowadzone badania potwierdziły możliwość budowy interpretowalnych klasyfikatorów dla danych temporalnych operujących bezpośrednio na danych czasowych, bez konieczności ich wcześniejszej agregacji, przy zachowaniu skuteczności porównywalnej z wybranymi metodami klasycznymi oraz metodami głębokiego uczenia.

Słowa kluczowe: dane temporalne, klasyfikacja danych, pojęcia czasowe, cięcia temporalne, drzewa decyzyjne, TC-Tree, VTC-Tree, eksploracja danych, interpretowalne modele uczenia maszynowego.