

Teoria uczenia dla modelowania różnicowego

Szymon Jaroszewicz
Instytut Podstaw Informatyki PAN, Warszawa
sj@ipipan.waw.pl

1. Opis projektu

Modelowanie różnicowe jest dziedziną uczenia maszynowego zajmującą się przewidywaniem wpływu danego działania (terapii medycznej, kampanii marketingowej) na konkretne osoby. Budowa modelu różnicowego korzysta z dwóch zbiorów uczących: eksperymentalnego (osoby poddane działaniu) i kontrolnego (osoby nie poddane działaniu). Zadanie polega na budowie modelu, który przewiduje różnice między prawdopodobieństwami klas w grupie eksperymentalnej i kontrolnej. Mimo rosnącej popularności modelowania różnicowego, niewiele jest rezultatów teoretycznych dotyczących tego problemu.

Niniejsza propozycja dotyczy stworzenia podstaw teorii uczenia dla modelowania różnicowego. Przykładowe pytania, na które postaramy się odpowiedzieć to: jakie warunki muszą być spełnione aby uczenie było możliwe, jaką rolę odgrywa w modelowaniu różnicowym wymiar Vapnika-Czerwonenkisa, jak dużo próbek uczących potrzeba by zagwarantować zadaną jakość modelu. Zdobyta wiedza posłuży do opracowania nowych algorytmów modelowania różnicowego.

2. Wymagania (oczekiwania)

- a. Ukończone studia drugiego stopnia z informatyki, matematyki lub fizyki
- b. Dobra znajomość rachunku prawdopodobieństwa i statystyki,
- c. Umiejętność programowania
- d. Dobra znajomość języka angielskiego

Bibliografia

1. Szymon Jaroszewicz, Uplift Modeling. W: Encyclopedia of Machine Learning and Data Mining, Springer, 2016
2. Ł. Zaniewicz, S. Jaroszewicz, Lp-Support vector machines for uplift modeling, Knowledge and Information Systems, <http://link.springer.com/content/pdf/10.1007/s10115-017-1040-6.pdf>