



UNIWERSYTET
IM. ADAMA MICKIEWICZA
W POZNANIU

Interpretacyjne uczenie maszynowe

Sylabus zajęć

Informacje podstawowe

Kierunek studiów Analiza i przetwarzanie danych Specjalność - Jednostka organizacyjna Wydział Matematyki i Informatyki Poziom studiów studia drugiego stopnia Forma studiów studia stacjonarne Profil studiów profil ogólnoakademicki		Cykl dydaktyczny 2023/24 Kod zajęć 06APDS.28K.00342.23 Języki wykładowe polski Obligatoryjność Fakultatywny Blok zajęciowy Przedmioty kierunkowe
Koordynator zajęć	Jarosław Jasiewicz	
Prowadzący zajęcia	Jarosław Jasiewicz	
Okres Semestr 4	Forma zajęć / liczba godzin / forma zaliczenia • Laboratorium: 30, Zaliczenie z oceną	Liczba punktów ECTS 3

Cele kształcenia dla zajęć

Kod	Cel
C1	Prezentowanie metod wydobycia wiedzy z algorytmów uczenia maszynowego do rozwiązywania złożonych problemów analizy danych.

Wymagania wstępne

Znajomość metod uczenia maszynowego

Efekty uczenia się dla zajęć

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
Wiedzy - Student/ka:			
W1	rozumie znaczenie interpretowalności i wyjaśnialności modeli regresyjnych i klasyfikacyjnych w analizie danych.	APD_K2_W01, APD_K2_W03	Projekt
W2	zna strukturę bezpośrednio interpretowanych algorytmów uczenia maszynowego (CART, LS, GLM, GAM).	APD_K2_W01, APD_K2_W03	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
W3	zna metody budowania modeli wyjaśniających dla dowolnych algorytmów uczenia maszynowego.	APD_K2_W01, APD_K2_W03	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
W4	zna metody wydobywania istotnych informacji dla algorytmów działających na danych uporządkowanych (np. obrazach).	APD_K2_W01, APD_K2_W03	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
Umiejętności - Student/ka:			
U1	potrafi wydobywać z algorytmów uczących oraz interpretować cechy decydujące o wyniku działania algorytmu.	APD_K2_U01, APD_K2_U02	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
U2	potrafi stosować globalne modele wyjaśniające (poziom modelu) dla złożonych algorytmów uczenia maszynowego.	APD_K2_U01, APD_K2_U03, APD_K2_U05, APD_K2_U08	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
U3	potrafi stosować lokalne modele wyjaśniające (poziom przypadku) dla złożonych algorytmów uczenia maszynowego.	APD_K2_U01, APD_K2_U03, APD_K2_U05	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
U4	potrafi stosować narzędzia wizualne stosowane w interpretacyjnym uczeniu maszynowym.	APD_K2_U01, APD_K2_U03, APD_K2_U07, APD_K2_U08	Projekt, Zaliczenie praktyczne (analiza wykonawstwa)
Kompetencji społecznych - Student/ka:			
K1	jest gotów/gotowa wykorzystać interpretacyjne uczenie maszynowe w samodzielnie realizowanym projekcie analizy danych.	APD_K2_K03	Projekt
K2	jest gotów/gotowa do prezentacji wyników analiz w sposób przystępny dla odbiorcy.	APD_K2_K02, APD_K2_K06	Projekt

Treści programowe dla zajęć

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
1.	Podstawy interpretacyjnego uczenia maszynowego, podstawowe pojęcia, zastosowania,	W1	Laboratorium
2.	Algorytmy uczenia maszynowego o interpretowanej strukturze, metody wydobywania wiedzy, rola błędów w interpretacji algorytmów, ocena interpretowalności	W2, U1	Laboratorium

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
3.	Wydobywanie istotnych cech z podstawowych algorytmów uczenia maszynowego, wyszukiwanie przypadków odstających i ich wpływ na interpretację	W3, U1	Laboratorium
4.	Tworzenie modeli wyjaśniających dla algorytmów typu „czarna skrzynka”, interpretacja modeli oparta na przypadkach	W3	Laboratorium
5.	Globalne modele wyjaśniające, wykresy częściowej zależności, interakcja cech, dekompozycja funkcjonalna	W3, U2, K2	Laboratorium
6.	Lokalne modele wyjaśniające DeepLIFT, SHAP, LIME, DALEX	W3, U3, K2	Laboratorium
7.	Podstawowe metody wizualizacji algorytmów uczenia maszynowego i ich interpretacja	U4, K2	Laboratorium
8.	Wydobywanie cech dla algorytmów działających na danych uporządkowanych, transformacje danych w celu wydobycia nowych cech	W4	Laboratorium
9.	Projekt zaliczeniowy	W3, U1, U4, K1	Laboratorium

Informacje dodatkowe

Forma zajęć	Metody i formy prowadzenia zajęć
Laboratorium	Metoda analizy przypadków, Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych), Metoda projektu

Forma zajęć	Warunki zaliczenia zajęć
Laboratorium	Wypełnienie założeń projektu (50%) i rozwiązanych dwóch zadań (2* 25%). • Bardzo dobry (Bdb; 5,0) = od 90% punktów • Dobry plus (Db+; 4,5) = od 80% punktów • Dobry (Db; 4,0) = od 70% punktów • Dostateczny plus (Dst+; 3,5) = od 60% punktów • Dostateczny (Dst; 3,0) = od 50% punktów • Niedostateczny (Ndst; 2,0) = poniżej 50% punktów

Literatura

Obowiązkowa

1. Molnar, Christop, 2022, Interpretable Machine Learning, <https://christophm.github.io/interpretable-ml-book/>

Dodatkowa

1. Biecek, Przemysław, Burzykowski Tomasz, 2020, Explanatory Model Analysis, <https://ema.drwhy.ai/>
2. Xuhong, Li et al. 2021, Interpretable Deep Learning: Interpretation, Interpretability, Trustworthiness, and Beyond, <https://arxiv.org/pdf/2103.10689.pdf>

Nakład pracy studenta i punkty ECTS

Rodzaje zajęć studenta	Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć
Laboratorium	30
Przygotowanie do zajęć	20
Czytanie wskazanej literatury	10
Przygotowanie projektu	20
Łączny nakład pracy studenta	Liczba godzin 80
Liczba punktów ECTS	ECTS 3

* godzina (lekcyjna) oznacza 45 minut

Efekty uczenia się dla kierunku

Kod	Treść
APD_K2_K02	Absolwent/ka jest gotów/gotowa do zrozumienia i pogłębienia świadomości roli informatyki w kształtowaniu życia społecznego
APD_K2_K03	Absolwent/ka jest gotów/gotowa do zrozumienia ograniczenia własnej wiedzy i rozumie potrzebę dalszego kształcenia, w tym zdobywania wiedzy pozadziśniskowej
APD_K2_K06	Absolwent/ka jest gotów/gotowa do zrozumienia najważniejszych osiągnięć w swojej dziedzinie i stojących przed nią wyzwań; potrafi je przedstawić laikom w sposób popularny
APD_K2_U01	Absolwent/ka potrafi wykorzystywać zaawansowane metody, narzędzia i technologie informatyczne do rozwiązywania problemów teoretycznych i problemów praktycznych w wybranych obszarach zastosowań
APD_K2_U02	Absolwent/ka potrafi zastosować zaawansowaną wiedzę matematyczną do formułowania, analizowania i rozwiązywania złożonych zadań związanych z analizą i przetwarzaniem danych
APD_K2_U03	Absolwent/ka potrafi pisać, uruchamiać i testować złożone programy w wybranych środowiskach programistycznym
APD_K2_U05	Absolwent/ka potrafi projektować, analizować i stosować zaawansowane algorytmy
APD_K2_U07	Absolwent/ka potrafi posługiwać się różnymi formatami reprezentacji rozmaitych rodzajów danych, w szczególności integrować dane pochodzące z różnych źródeł, transformować je do wspólnej spójnej postaci
APD_K2_U08	Absolwent/ka potrafi projektować i implementować zaawansowane struktury danych
APD_K2_W01	Absolwent/ka zna i rozumie w pogłębionym stopniu współczesne metody, narzędzia i technologie informatyczne właściwe dla wybranych obszarów zastosowań niezbędne przy analizie i przetwarzaniu danych pochodzących z różnych źródeł
APD_K2_W03	Absolwent/ka zna i rozumie zasady rozwiązywania problemów z wykorzystaniem zaawansowanych algorytmów i struktur danych