Uczenie maszynowe

Sylabus zajęć

Informacje podstawowe

Kierunek studiów Informatyka	**Cykl dydaktyczny** 2023/24
Specjalność -	**Kod zajęć** 06INFN.310S.02038.23
Jednostka organizacyjna Wydział Matematyki i Informatyki	**Języki wykładowe** polski
Poziom studiów studia inżynierskie pierwszego stopnia	**Obligatoryjność** Fakultatywny
Forma studiów studia niestacjonarne	**Blok zajęciowy** Przedmioty specjalnościowe
Profil studiów profil ogólnoakademicki	

Koordynator zajęć	Paweł Skórzewski
Prowadzący zajęcia	Paweł Skórzewski

Okres	**Forma zajęć / liczba godzin / forma zaliczenia**	**Liczba punktów ECTS**
Semestr 5	• Wykład: 15, Egzamin • Laboratorium: 15, Zaliczenie z oceną	5

Cele kształcenia dla zajęć

Kod	Cel
C1	wyjaśnienie koncepcji uczenia maszynowego
C2	zaprezentowanie najważniejszych algorytmów uczenia maszynowego
C3	wykształcenie u studentów umiejętności stosowania metod uczenia maszynowego w praktyce programistycznej
C4	wykształcenie u studentów umiejętności poprawnej ewaluacji rozwiązań programistycznych wykorzystujących metody uczenia maszynowego

Wymagania wstępne

- Podstawowa umiejętność programowania.
- Znajomość podstaw algebry liniowej.

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	zna i rozumie rolę, znaczenie i zastosowania uczenia maszynowego we współczesnej informatyce	INF_K3_W09	Egzamin pisemny
W2	zna podstawowe typy zadań uczenia maszynowego i ich przykłady	INF_K3_W09	Egzamin pisemny
W3	zna i rozumie zagadnienie regresji liniowej jednej i wielu zmiennych	INF_K3_W09	Egzamin pisemny
W4	zna i rozumie metodę gradientu prostego	INF_K3_W09	Egzamin pisemny
W5	zna i rozumie zagadnienie regresji logistycznej	INF_K3_W09	Egzamin pisemny
W6	zna i rozumie metody i znaczenie ewaluacji algorytmów uczenia maszynowego	INF_K3_W09	Egzamin pisemny
W7	zna i rozumie rolę zbiorów danych: uczącego, walidacyjnego i testowego	INF_K3_W09	Egzamin pisemny
W8	zna i rozumie podstawowe miary jakości stosowane przy ewaluacji algorytmów uczenia maszynowego	INF_K3_W09	Egzamin pisemny
W9	zna i rozumie zjawiska nadmiernego i niedostatecznego dopasowania	INF_K3_W09	Egzamin pisemny
W10	zna i rozumie metody regularyzacji	INF_K3_W09	Egzamin pisemny
W11	zna i rozumie znaczenie optymalizacji i jej podstawowe metody	INF_K3_W09	Egzamin pisemny
W12	zna i rozumie ideę i najważniejsze algorytmy uczenia nienadzorowanego	INF_K3_W09	Egzamin pisemny
W13	zna i rozumie zasadę działania naiwnego klasyfikatora bayesowskiego	INF_K3_W09	Egzamin pisemny
W14	zna i rozumie zasadę działania algorytmu k najbliższych sąsiadów	INF_K3_W09	Egzamin pisemny
W15	zna i rozumie zasadę działania drzew decyzyjnych	INF_K3_W09	Egzamin pisemny
W16	zna i rozumie zasadę działania maszyn wektorów nośnych	INF_K3_W09	Egzamin pisemny
W17	zna i rozumie zasadę działania sztucznych sieci neuronowych, w tym wielowarstwowych	INF_K3_W09	Egzamin pisemny
W18	zna i rozumie ideę uczenia głębokiego	INF_K3_W09	Egzamin pisemny
W19	zna i rozumie zasadę działania i zastosowania splotowych sieci neuronowych	INF_K3_W09	Egzamin pisemny
W20	zna i rozumie zasadę działania i zastosowania rekurencyjnych sieci neuronowych	INF_K3_W09	Egzamin pisemny

Kod	Efekty uczenia się dla zajęć w zakresie	Efekty uczenia się dla kierunku	Metody weryfikacji osiągnięcia efektów uczenia się dla zajęć
W21	zna i rozumie zasadę działania i zastosowania modeli typu encoder-decoder, w szczególności do tworzenia modeli języka i tłumaczenia maszynowego	INF_K3_W09	Egzamin pisemny
W22	zna i rozumie ideę oraz podstawowe paradygmaty uczenia przez wzmacnianie	INF_K3_W09	Egzamin pisemny
Umiejętności – Student/ka:			
U1	umie korzystać z podstawowych narzędzi bibliotek NumPy i PyTorch oraz elementów języka Python przydatnych do implementowania rozwiązań z dziedziny uczenia maszynowego	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U2	umie przetwarzać dane przechowywane w tekstowych formatach tabelarycznych (CSV/TSV)	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U3	umie wizualizować dane, korzystając z bibliotek Matplotlib i Seaborn	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U4	umie zaimplementować algorytm gradientu prostego do znalezienia rozwiązania problemu regresji liniowej	INF_K3_U05_inz, INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U5	umie zaimplementować algorytm gradientu prostego do znalezienia rozwiązania problemu regresji logistycznej	INF_K3_U05_inz, INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U6	umie dokonać odpowiedniego podziału danych na zbiory: uczący, walidacyjny i testowy, i wykorzystać je do ewaluacji rozwiązania zadania z dziedziny uczenia maszynowego	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U7	umie korzystać z modułów pakietu Scikit-Learn do implementacji rozwiązań uczenia maszynowego	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U8	umie zapobiegać nadmiernemu i niedostatecznemu dopasowaniu w implementowanych przez siebie rozwiązaniach	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U9	umie poprawnie reprezentować dane różnych typów i korzystać z nich do rozwiązywania problemów metodami uczenia maszynowego	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U10	umie stosować metody optymalizacji uczenia maszynowego	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U11	umie zaimplementować przykładowy algorytm uczenia nienadzorowanego	INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U12	umie wykorzystać metodę propagacji wstecznej do uczenia wielowarstwowych sieci neuronowych	INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U13	umie zaimplementować sieć neuronową z wykorzystaniem odpowiednich bibliotek	INF_K3_U06_inz, INF_K3_U07_inz	Zadania wykonywane podczas zajęć
U14	umie zaprojektować, zaimplementować i zewaluować system wykorzystujący uczenie maszynowe	INF_K3_U03, INF_K3_U05_inz, INF_K3_U06_inz, INF_K3_U07_inz, INF_K3_U09	Projekt

Treści programowe dla zajęć

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
1.	Wprowadzenie do uczenia maszynowego. Czym jest uczenie maszynowe? Rola i miejsce uczenia maszynowego we współczesnej informatyce. Przegląd zastosowań i metod uczenia maszynowego. Podstawowe pojęcia związane z uczeniem maszynowym.	W1, W2	Wykład
2.	Podstawowe narzędzia uczenia maszynowego. Elementy języka Python przydatne przy implementowaniu algorytmów uczenia maszynowego. Biblioteki NumPy i PyTorch	U1	Laboratorium
3.	Narzędzia przetwarzania i wizualizacji danych w języku Python. Format CSV/TSV. Biblioteki Matplotlib i Seaborn.	U2, U3	Laboratorium
4.	Regresja liniowa jednej zmiennej. Funkcja kosztu. Metoda gradientu prostego. Regresja liniowa wielu zmiennych.	W3, W4	Wykład
5.	Implementacja regresji liniowej jednej zmiennej w języku Python.	U4	Laboratorium
6.	Regresja logistyczna. Metoda gradientu prostego dla regresji logistycznej.	W5	Wykład
7.	Implementacja regresji logistycznej w języku Python.	U5	Laboratorium
8.	Ewaluacja algorytmów uczenia maszynowego. Podział na zbiory: uczący, testowy i walidacyjny. Walidacja krzyżowa. Miary jakości.	W6, W7, W8	Wykład
9.	Pakiet Scikit-Learn. Implementacja regresji liniowej i regresji logistycznej z wykorzystaniem gotowych modułów. Implementacja wybranych metod ewaluacji.	U6, U7	Laboratorium
10.	Sposoby reprezentacji danych. Implementacja algorytmów regresji dla danych różnych typów, w tym dla danych nieliczbowych, oraz dla danych niepełnych.	U9	Laboratorium
11.	Nadmierne i niedostateczne dopasowanie. Obciążenie i wariancja. Ilustracja problemu nadmiernego dopasowania na przykładzie regresji wielomianowej. Metody regularyzacji.	W10, W9	Wykład
12.	Nadmierne i niedostateczne dopasowanie w praktyce. Implementacja metod zapobiegających nadmiernemu dopasowaniu.	U8	Laboratorium
13.	Stochastic Gradient Descent. Przegląd metod optymalizacji.	W11	Wykład
14.	Porównanie różnych metod optymalizacji na przykładach.	U10	Laboratorium
15.	Uczenie nienadzorowane. Algorytm k średnich. Algorytm analizy głównych składowych.	W12	Wykład
16.	Implementacja metod uczenia nienadzorowanego na przykładzie algorytmu k średnich.	U11	Laboratorium

Lp.	Treści programowe dla zajęć	Efekty uczenia się dla zajęć	Formy zajęć
17.	Przegląd metod uczenia nadzorowanego. Naiwny klasyfikator bayesowski. Algorytm k najbliższych sąsiadów. Drzewa decyzyjne. Maszyny wektorów nośnych.	W13, W14, W15, W16	Wykład
18.	Wprowadzenie do sztucznych sieci neuronowych. Prosty perceptron. Funkcje aktywacji. Wielowarstwowe sieci neuronowe. Propagacja wsteczna. Uczenie wielowarstwowych sieci neuronowych.	W17	Wykład
19.	Implementacja sieci neuronowych.	U12, U13	Laboratorium
20.	Splotowe sieci neuronowe – idea, przegląd najpopularniejszych architektur, przegląd zastosowań. Czym jest uczenie głębokie?	W18, W19	Wykład
21.	Rekurencyjne sieci neuronowe – idea, przegląd najpopularniejszych architektur, przegląd zastosowań. Modele typu encoder-decoder. Neuronowe tłumaczenie maszynowe. Autoencoder. Word embeddings.	W20, W21	Wykład
22.	Wprowadzenie do uczenia przez wzmacnianie.	W22	Wykład
23.	Indywidualny projekt programistyczny – implementacja wybranych metod uczenia maszynowego.	U14	Laboratorium

Informacje dodatkowe

Forma zajęć	Metody i formy prowadzenia zajęć
Wykład	Wykład z prezentacją multimedialną wybranych zagadnień
Laboratorium	Rozwiązywanie zadań (np.: obliczeniowych, artystycznych, praktycznych), Metoda laboratoryjna

Forma zajęć	Warunki zaliczenia zajęć
Wykład	Na ocenę końcową składa się wynik uzyskany na egzaminie pisemnym. Egzamin ma formę testu jednokrotnego wyboru. Test składa się z 50 pytań. Za poprawną odpowiedź na każde pytanie student otrzymuje 1 punkt. O ocenie będzie decydować suma uzyskanych punktów, zgodnie z poniższym przelicznikiem (skala ocen): • od 26 pkt. - ocena dostateczna (3) • od 31 pkt. - ocena dostateczna+ (3,5) • od 36 pkt. - ocena dobra (4) • od 41 pkt. - ocena dobra+ (4,5) • od 46 pkt. - ocena bardzo dobra (5)
Laboratorium	Punkty można zdobyć za: • zadania na zajęciach (max. 66 pkt.): ◦ po 6 pkt. do zdobycia na każdych z 11 pierwszych zajęć • projekt programistyczny (max. 34 pkt.) Ocena z laboratoriów zależy od sumy uzyskanych punktów (skala ocen): • od 51 pkt. - ocena dostateczna (3) • od 61 pkt. - ocena dostateczna+ (3,5) • od 71 pkt. - ocena dobra (4) • od 81 pkt. - ocena dobra+ (4,5) • od 91 pkt. - ocena bardzo dobra (5)

Literatura

Obowiązkowa

1. S. Raschka, Python Machine Learning, Packt, Birmingham 2015
2. S. Marsland, Machine Learning: An Algorithmic Perspective, CRC, Boca Raton 2015
3. W. Richert, L.P. Coelho, Building Machine Learning Systems with Python, Packt, Birmingham 2013
4. G. Moncecchi, R. Garreta, Learning scikit-learn: Machine Learning in Python, Packt, Birmingham 2013
5. K. Krawiec, J. Stefanowski, Uczenie maszynowe i sieci neuronowe, WPP, Poznań 2004
6. M. Krzyśko, W. Wołyński, T. Górecki, M. Skorzybut, Systemy uczące się, WNT, Warszawa 2008
7. W. Duch, J. Korbicz, L. Rutkowski, R. Tadeusiewicz, Sieci neuronowe, Exit, Warszawa 2000
8. K.P. Murphy, Machine Learning: a Probabilistic Perspective, 2015

Dodatkowa

1. M. Nielsen, Neural Networks and Deep Learning, neuralnetworksanddeeplearning.com (odczyt: 2020-11-03)
2. J. Brownlee, Machine Learning Mastery, machinelearningmastery.com com (odczyt: 2020-11-03)
3. I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, www.deeplearningbook.org com (odczyt: 2020-11-03)
4. A. Ng, Machine Learning, Coursera – kurs online, https://www.coursera.org/learn/machine-learning (odczyt: 2020-11-03)
5. G. Hinton, Neural Networks for Machine Learning, Coursera – kurs online, https://www.coursera.org/learn/neural-networks (odczyt: 2020-11-03)
6. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Nakład pracy studenta i punkty ECTS

Rodzaje zajęć studenta	**Średnia liczba godzin* przeznaczonych na zrealizowane rodzaje zajęć**
Wykład	15
Laboratorium	15
Przygotowanie do zajęć	20
Czytanie wskazanej literatury	20
Przygotowanie projektu	40
Inne	15
Łączny nakład pracy studenta	**Liczba godzin** 125
Liczba punktów ECTS	**ECTS** 5

* godzina (lekcyjna) oznacza 45 minut

Efekty uczenia się dla kierunku

Kod	Treść
INF_K3_U03	Absolwent/ka potrafi pracować indywidualnie i w zespole, zarządzać swoim czasem oraz podejmować zobowiązania i dotrzymywać terminów
INF_K3_U05_inz	Absolwent/ka potrafi pisać, uruchamiać i testować programy w wybranym środowisku programistycznym
INF_K3_U06_inz	Absolwent/ka potrafi ocenić, na podstawowym poziomie, przydatność metod i narzędzi informatycznych oraz wybrać i zastosować właściwą metodę i narzędzia do typowych zadań informatycznych
INF_K3_U07_inz	Absolwent/ka potrafi zastosować wybrane metody wykorzystywane w wiodących kierunkach badań informatyki
INF_K3_U09	Absolwent/ka potrafi przygotować dokumentacje, opracowania i raporty w języku polskim i języku angielskim, w tym z wykorzystaniem podstawowych ujęć teoretycznych, a także różnych źródeł
INF_K3_W09	Absolwent/ka zna i rozumie zaawansowane zagadnienia związane ze sztuczną inteligencją