



Machine learning

Course description sheet

Basic information

Field of study Automatics and Robotics		Didactic cycle 2021/2022
Major -		Course code EAiRS.li200.76f98a4ab46ffd16f2e0e9a2bfac8cb8.21
Organisational unit Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering		Lecture languages polish
Study level First-cycle (engineer) programme		Mandatoriness Elective
Form of study Full-time studies		Block General Modules
Profile General academic		Course related to scientific research Yes
Course coordinator	Joanna Jaworek-Korjakowska	
Lecturer	Joanna Jaworek-Korjakowska, Anna Wójcicka, Dariusz Kucharski	
Period Semester 6	Method of verification of the learning outcomes Completing the classes Activities and hours Lectures: 28 Laboratory classes: 28	Number of ECTS credits 4

Goals

C1	Zapoznanie Studenta z podstawowymi i zaawansowanymi metodami uczenia maszynowego, zasadami ich działania oraz możliwościami zastosowania.
C2	Przekazanie wiedzy z zakresu narzędzi i środowisk do tworzenia i rozbudowy systemów informatycznych wykorzystujących algorytmy uczenia maszynowego
C3	Uświadomienie słuchaczom potrzeby wykorzystania zaawansowanych metod uczenia maszynowego tak, aby w sposób samodzielny i kreatywny byli w stanie wybrać odpowiednie rozwiązanie dla określonego problemu badawczego.

Course's learning outcomes

Code	Outcomes in terms of	Learning outcomes prescribed to a field of study	Methods of verification
Knowledge - Student knows and understands:			
W1	Zna podstawowe i zaawansowane metody uczenia maszynowego, zasady ich działania oraz możliwości zastosowania.	AiR1A_W01	Test, Report
W2	Zna narzędzia i środowiska do tworzenia i rozbudowy systemów informatycznych wykorzystujących algorytmy uczenia maszynowego	AiR1A_W04	Test, Report
Skills - Student can:			
U1	Potrafi pozyskiwać informacje o zaawansowanych metodach uczenia maszynowego oraz wykorzystywać je podczas implementowania rozwiązań algorytmicznych.	AiR1A_U01	Activity during classes, Report
U2	Potrafi przygotować dokumentację zaimplementowanego rozwiązania ze szczegółowym omówieniem wyników, wyciągnąć wnioski oraz wyczerpująco je uzasadnić.	AiR1A_U03	Activity during classes, Report
Social competences - Student is ready to:			
K1	Potrafi zaproponować ulepszenie oraz optymalizację zaimplementowanego algorytmu uczenia maszynowego. Potrafi przeanalizować wyniki i na ich podstawie sformułować wnioski.	AiR1A_K01	Activity during classes, Report
K2	Zna i rozumie potrzebę wykorzystania zaawansowanych metod uczenia maszynowego. Potrafi w sposób samodzielny i kreatywny wybrać odpowiednie rozwiązanie dla określonego problemu badawczego.	AiR1A_K01	Activity during classes, Report

Program content ensuring the achievement of the learning outcomes prescribed to the module

Podczas zajęć omówione zostaną zarówno podstawowe jak i zaawansowane metody uczenia maszynowego i sztucznej inteligencji w tym sieci neuronowych. Studenci zostaną przygotowani do praktycznego wykorzystania najnowszych algorytmów i rozwiązań.

Student workload

Activity form	Average amount of hours* needed to complete each activity form
Lectures	28
Laboratory classes	28
Realization of independently performed tasks	44
Student workload	Hours 100
Workload involving teacher	Hours 56

* hour means 45 minutes

Program content

No.	Program content	Course's learning outcomes	Activities
-----	-----------------	----------------------------	------------

1.	<p>1. Wprowadzenie do tematyki uczenia maszynowego, podstawowe zagadnienia z algebry</p> <p>2. Wykorzystywanie środowiska Python do uczenia maszynowego. Regresja liniowa wielu zmiennych (optymalizacja danych).</p> <p>3. Zagadnienie klasyfikacji: algorytm K-najbliższych sąsiadów (k-NN), okno Parzena, drzewa decyzyjne</p> <p>4. Sprawdzanie poprawności danych. Kwestia brakujących danych. Dobór odpowiednich cech. Ocenianie istotności cech za pomocą algorytmu losowego lasu. Analiza wyników klasyfikacji (macierz pomyłek, krzywa ROC)</p> <p>5. Regularyzacja (problem nadmiernego dopasowania, funkcja kosztu)</p> <p>6. Strategia tworzenia systemów uczenia maszynowego</p> <p>7. Maszyny wektorów nośnych (ang. Support Vector Machine - SVM) - teoria, funkcje jądra, wyznaczanie maksymalnego marginesu za pomocą maszyn wektorów nośnych.</p> <p>8. Redukcja wymiarowości: analiza głównych składowych (PCA) - motywacja, sformułowanie problemu, algorytm</p> <p>9. Detekcja anomalii w dużych zbiorach danych (ang. Large Scale Machine Learning - LSML)</p> <p>10. Najlepsze metody oceny modelu i strojenie parametryczne.</p> <p>11. Wprowadzenie do sztucznych sieci neuronowych. Podstawowe zagadnienia i domeny sztucznej inteligencji. Sztuczny neuron. Wielowarstwowe sieci neuronowe I</p> <p>12. Sieci wielowarstwowe II. Algorytm backpropagation, i jego modyfikacje.</p> <p>13. Sposoby uczenia sieci neuronowych (z nauczycielem i bez nauczyciela). Omówienie sieci neuronowych typu RBF, sieci samoorganizujące się, sieci Kohonena.</p> <p>14. Omówienie sieci LVQ, rezonansowej ART, Hopfielda. Praktyczne zastosowanie metod uczenia maszynowego.</p>	W1, W2, K2	Lectures
----	--	------------	----------

2.	<p>1. Wprowadzenie do programowania w Pythonie z wykorzystaniem pakietów Panda, Scikit-learn, NumPy.</p> <p>2. Regresja liniowa wielu zmiennych - (optymalizacja danych).</p> <p>3. Zagadnienie klasyfikacji: algorytm K-najbliższych sąsiadów (k-NN), drzewa decyzyjne. Wybór liczby sąsiadów/rozmiaru okna, wpływ obserwacji odstających.</p> <p>4. Tworzenie i analiza drzew decyzyjnych w wersji tekstowej i graficznej. Analiza procesu klasyfikacji, kosztu resubstytucji (ang. resubstitution error). Analiza wyników klasyfikacji (macierz pomyłek, krzywa ROC)</p> <p>5. Regularyzacja (problem nadmiernego dopasowania, funkcja kosztu). Analiza wyników klasyfikacji (macierz pomyłek, krzywa ROC)</p> <p>6. Walidacja: różnice w wynikach między poprawnym a błędnym zastosowaniem. Przygotowanie zbioru danych: strategia postępowania w przypadku brakujących danych (ang. data imputation), analiza błędnych pomiarów (ang. outlier detection), wpływ normalizacji danych na wynik klasyfikatorów opartych o miary odległości. Ocenianie istotności cech za pomocą algorytmu losowego lasu.</p> <p>7. Klasyfikacja parametrów geometrycznych znamion skórnych przy pomocy SVM</p> <p>8. Kompresja zdjęć twarzy metodą PCA. Omówienie i wizualizacja wyników.</p> <p>9. Wykorzystanie wielowymiarowego rozkładu Gaussa do detekcji anomalii w danych biomedycznych (analiza częstości skurczów serca). Zastosowanie algorytmu expectation-minimization do maksymalizacji funkcji wiarygodności.</p> <p>10. Usprawnianie cyklu pracy za pomocą kolejowania. Stosowanie k-krotnego sprawdzianu krzyżowego w ocenie skuteczności modelu. Sprawdzanie algorytmów za pomocą krzywych uczenia i krzywych walidacji</p> <p>11. Implementacja podstawowej komórki neuronowej. Sieci neuronowe typu Perceptron prosty. Architektura sieci MLP</p> <p>12. Przygotowanie i zastosowanie jednowarstwowej perceptronowej sieci neuronowej do klasyfikacji obrazów i praktyczne jej wykorzystanie</p> <p>13. Kompresja obrazów wykonana przy pomocy samoorganizujących się map. Zastosowanie sieci Kohonena do rozpoznawania obiektów.</p> <p>14. Zastosowanie sieci LVQ do problemu rozpoznawania 3 klas. Sieć rezonansowa ART-rozpoznawanie liter.</p>	W1, W2, U1, U2, K1, K2	Laboratory classes
----	--	------------------------	--------------------

Extended information/Additional elements

Teaching methods and techniques:

Lectures, E-learning

Activities	Methods of verification	Credit conditions
Lectures	Test	Wykład: – Obecność obowiązkowa: Nie

Activities	Methods of verification	Credit conditions
Lab. classes	Activity during classes, Test, Report	Laboratorium: 1. Warunkiem uczestnictwa w zajęciach laboratoryjnych jest dokonanie zapisu na kurs e-learningowy na platformie AGH. 2. Obecność na zajęciach jest obowiązkowa, dopuszczalna jest jedna nieusprawiedliwiona nieobecność. Laboratoria zaległe należy odrobić w ciągu tygodnia od powrotu na uczelnię po nieobecności, w trakcie zajęć pozostałych grup na roku lub podczas konsultacji. Nieodrobienie zajęć (brak sprawozdania) w tym okresie skutkować będzie definitywnym brakiem zaliczenia danego ćwiczenia. Dopuszczalne jest jedno niezaliczone ćwiczenie. Kolokwium: W trakcie semestru odbędą się dwa kolokwia zaliczeniowe (odpowiednio w połowie oraz pod koniec semestru). Wymagane jest uzyskanie powyżej 50 % punktów z każdego z kolokwium. Studentom przysługuje możliwość jednokrotnej poprawy danego kolokwium.

Conditions and the manner of completing each form of classes, including the rules of making retakes, as well as the conditions for admission to the exam

Laboratorium: Obecność na zajęciach jest obowiązkowa, dopuszczalna jest jedna nieusprawiedliwiona nieobecność. Laboratoria zaległe należy odrobić w ciągu tygodnia od powrotu na uczelnię po nieobecności, w trakcie zajęć pozostałych grup na roku lub podczas konsultacji. Nieodrobienie zajęć (brak sprawozdania) w tym okresie skutkować będzie definitywnym brakiem zaliczenia danego ćwiczenia. Dopuszczalne jest jedno niezaliczone ćwiczenie. Kolokwium: W trakcie semestru odbędą się dwa kolokwia zaliczeniowe (odpowiednio w połowie oraz pod koniec semestru). Wymagane jest uzyskanie powyżej 50 % punktów z każdego z kolokwium. Studentom przysługuje możliwość jednokrotnej poprawy danego kolokwium. W przypadku poprawy kolokwium student może uzyskać maksymalnie ocenę 3.0 (zal.) Ocena końcowa: Podstawą do wystawienia oceny z laboratorium jest średnia arytmetyczna ocen uzyskanych za każde z kolokwium zaokrąglona w górę do najbliższej oceny zgodnej z Regulaminem Studiów AGH. Aby zaliczyć przedmiot należy uzyskać zaliczenie z laboratoriów oraz z każdego z kolokwium. Uwaga. Dla osób zaangażowanych w działalność kół naukowych (projekty, konkursy) lub w prace badawcze istnieje możliwość realizacji zadań ustalonych z prowadzącym przedmiot oraz opiekunem naukowym zamiast wybranych ćwiczeń laboratoryjnych. Zadania te muszą być zbieżne z tematyką przedmiotu i zapewniać realizację efektów kształcenia. Sposób ich wykonania, sporządzony raport oraz dostarczony kod źródłowy będą podstawą do wystawienia oceny z laboratorium.

Method of determining the final grade

Ocena końcowa (OK) jest wyznaczana na podstawie ocen z kolokwium zaliczeniowych (K1, K2) zgodnie ze wzorem: $OK = 0.5 * K1 + 0.5 * K2$

Manner and mode of making up for the backlog caused by a student justified absence from classes

Obecność na zajęciach jest obowiązkowa, dopuszczalna jest jedna nieusprawiedliwiona nieobecność. Laboratoria zaległe należy odrobić w ciągu tygodnia od powrotu na uczelnię po nieobecności, w trakcie zajęć pozostałych grup na roku lub podczas konsultacji. Nieodrobienie zajęć (brak sprawozdania) w tym okresie skutkować będzie definitywnym brakiem zaliczenia danego ćwiczenia. Dopuszczalne jest jedno niezaliczone ćwiczenie.

Prerequisites and additional requirements

Wymagania wstępne: zaliczenie z przedmiotu - Analiza i bazy danych

Wymagania dodatkowe: podstawowa znajomość Pythona oraz podstawowa wiedza ze statystyki

Rules of participation in given classes, indicating whether student presence at the lecture is obligatory

Lectures: Studenci uczestniczą w zajęciach poznając kolejne treści nauczania zgodnie z sylabusem przedmiotu. Studenci winni na bieżąco zadawać pytania i wyjaśniać wątpliwości. Rejestracja audiowizualna wykładu wymaga zgody prowadzącego. Laboratory classes: Studenci wykonują ćwiczenia laboratoryjne zgodnie z materiałami udostępnionymi przez prowadzącego. Student jest zobowiązany do przygotowania się w przedmiocie wykonywanego ćwiczenia, co może zostać zweryfikowane kolokwium w formie ustnej lub pisemnej. Zaliczenie zajęć odbywa się na podstawie zaprezentowania rozwiązania

postawionego problemu.

Literature

Obligatory

1. Cichosz P., Systemy uczące się, WNT Warszawa, 2000, ISBN 83-204-2544-1
2. Ian H. Witten, Eibe Frank, Mark A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 2011
3. Peter Flach, Machine Learning: The Art and Science of Algorithms that Make Sense of Data, 2012
4. Krawiec K., Stefanowski J.: Uczenie maszynowe i sieci neuronowe, Wydawnictwo Politechniki Poznańskiej, Poznań, 2004
5. Christopher M. Bishop: Pattern Recognition and Machine Learning, wyd. Springer, 2006
6. Richard O. Duda, Peter E. Hart, David G. Stork: Pattern Classification (2nd Ed.), wyd. Wiley, 2000

Scientific research and publications

Publications

1. J. Jaworek-Korjakowska: Computer-aided diagnosis of micro-malignant melanoma lesions applying support vector machines, BioMed Research International, 2016, s. 1-8
2. J. Jaworek-Korjakowska, P. Kłeczek: Automatic classification of specific melanocytic lesions using artificial intelligence, BioMed Research International, 2016, s. 1-17
3. J. Jaworek-Korjakowska: Artificial neural networks in the diagnosis of pigmented skin lesions: a review., Bio-Algorithms and Med-Systems, 2015 vol. 11 iss. 2, s. 36
4. Jaworek-Korjakowska J.: A deep learning approach to vascular structure segmentation in dermoscopy colour images, BioMed Research International, vol. 2018, pp. 1 - 8, 2018
5. Jaworek-Korjakowska J., Kłeczek P.: Region adjacency graph approach for acral melanocytic lesion segmentation, Applied Sciences (Basel), vol. 8, pp. 1 - 14, 2018

Learning outcomes prescribed to a field of study

Code	Content
AiR1A_K01	krytycznej oceny posiadanej wiedzy i odbieranych treści; uznawania znaczenia wiedzy w rozwiązywaniu problemów poznawczych i praktycznych oraz zasięgania opinii ekspertów w przypadku trudności z samodzielnym rozwiązaniem problemu
AiR1A_U01	wykorzystywać posiadaną wiedzę – formułować i rozwiązywać złożone i nietypowe problemy oraz wykonywać zadania w warunkach nie w pełni przewidywalnych przez: - właściwy dobór źródeł i informacji z nich pochodzących, dokonywanie oceny, krytycznej analizy i syntezy tych informacji, - dobór oraz stosowanie właściwych metod i narzędzi, w tym zaawansowanych technik informacyjno-komunikacyjnych
AiR1A_U03	planować i organizować pracę indywidualną oraz w zespole; współdziałać z innymi osobami w ramach prac zespołowych (także o charakterze interdyscyplinarnym)
AiR1A_W01	zaawansowane zagadnienia w zakresie matematyki, obejmującą algebrę ze szczególnym uwzględnieniem teorii macierzy, analizę, równania różniczkowe i probabilistykę z elementami statystyki w sposób pozwalający na rozwiązanie typowych zagadnień automatyki; oraz zagadnienia w zakresie fizyki, obejmującą mechanikę, termodynamikę, optykę, elektryczność i magnetyzm, fizykę jądrową oraz fizykę ciała stałego, w tym wiedzę niezbędną do zrozumienia podstawowych zjawisk fizycznych występujących w typowych systemach dynamicznych oraz w ich otoczeniu.
AiR1A_W04	podstawy programowania obiektowego i strukturalnego wraz z elementami inżynierii oprogramowania; metody realizacji obliczeń i optymalizacji w środowiskach informatycznych oraz wizualizacji tych wyników; sposoby realizacji projektów informatycznych; przetwarzanie informatyczne informacji, ze szczególnym uwzględnieniem danych wizyjnych; oraz metodykę tworzenia ciągłych i dyskretnych modeli matematycznych.