

Sylabus przedmiotu / modułu kształcenia		
Nazwa przedmiotu/modułu kształcenia:		Uczenie maszynowe
Nazwa w języku angielskim:	Machine learning	
Język wykładowy:	polski	
Kierunek studiów, dla którego przedmiot jest oferowany:		informatyka
Jednostka realizująca:	Wydział Nauk Ścisłych i Przyrodniczych	
Rodzaj przedmiotu/modułu kształcenia (obowiązkowy/fakultatywny):		fakultatywny
Poziom modułu kształcenia (np. pierwszego lub drugiego stopnia):		pierwszego stopnia
Rok studiów:	trzeci	
Semestr:	szósty	
Liczba punktów ECTS:	4	
Imię i nazwisko koordynatora przedmiotu:		dr Piotr Świtalski
Imię i nazwisko prowadzących zajęcia:		mgr Mateusz Przychodzki
Założenia i cele przedmiotu:		Celem przedmiotu jest wprowadzenie studentów do metod uczenia maszynowego. Na wykładzie studenci zapoznają się z typowymi technikami klasyfikacji danych, metod regresji oraz redukcji wielowymiarowości. Omówione zostaną podstawy sztucznych sieci neuronowych. Zajęcia laboratoryjne poświęcone będą implementacji i weryfikacji metod oraz algorytmów poznanych na wykładzie.
Symbol efektu	Efekt uczenia się: WIEDZA	Symbol efektu kierunkowego
W_01	Zna i rozumie najważniejsze paradygmaty i metody problemu uczenia maszynowego	K_W01, K_W10
W_02	Ma wiedzę na temat algorytmów nadzorowanego i nienadzorowanego uczenia maszynowego	K_W10
W_03	Wie jak ocenić i poprawić jakość modeli uczenia maszynowego	K_W01, K_W10
Symbol efektu	Efekt uczenia się: UMIEJĘTNOŚCI	Symbol efektu kierunkowego
U_01	Potrafi wykorzystać biblioteki implementujące algorytmy uczenia maszynowego do rozwiązania postawionego problemu	K_U02, K_U22
U_02	Potrafi zaprojektować i przygotować implementację danego algorytmu uczenia maszynowego, odpowiednią dla problemu i dopasowaną do posiadanych danych	K_U07, K_U09

Forma i typy zajęć:	Studia stacjonarne: wykłady (21 godz.), ćwiczenia laboratoryjne (24 godz.) Studia niestacjonarne: wykłady (15 godz.), ćwiczenia laboratoryjne (15 godz.)
Wymagania wstępne i dodatkowe:	
Warunkiem uczestnictwa w zajęciach jest znajomość zagadnień z przedmiotów „Analiza danych” oraz „Sztuczna inteligencja” lub znajomość literatury obowiązującej w tych przedmiotach.	
Treści modułu kształcenia:	
Wprowadzenie do uczenia maszynowego. Rodzaje systemów uczenia maszynowego. Uczenie nadzorowane i uczenie nienadzorowane. Główne problemy uczenia maszynowego. Ocena przydatności danych. Dane źródłowe. Zbiory danych uczące i testowe. Korelacja. Reprezentatywność danych. Duplikaty. Szeregi czasowe. Przetwarzanie danych. Poprawianie błędnych danych. Klasyfikacja. Uczenie klasyfikatora binarnego. Miary wydajności. Klasyfikacja wieloklasowa. Analiza błędów. Klasyfikacja wieloetykietowa. Klasyfikacja wielowyjściowa. Klasyfikacja probabilistyczna. Drzewa decyzyjne. Uczenie i wizualizowanie drzewa decyzyjnego. Wyliczanie prognoz. Szacowanie prawdopodobieństw przynależności do klas. Algorytm uczący CART. Złożoność obliczeniowa. Hiperparametry regularyzacyjne. Regresja. Niestabilność. Maszyny wektorów nośnych. Liniowa klasyfikacja SVM. Klasyfikacja miękkiego marginesu. Nieliniowa klasyfikacja SVM. Jądro wielomianowe. Cechy podobieństwa. Gaussowskie jądro RBF. Złożoność obliczeniowa. Regresja SVM. Funkcja decyzyjna i prognozy. Programowanie kwadratowe. Problem dualny. Kernelizowane maszyny SVM. Przyrostowe maszyny SVM. Regresja. Model regresji liniowej. Gradient prosty. Regresja wielomianowa. Krzywe uczenia. Regularyzowane modele liniowe. Regresja logistyczna. Redukcja wielowymiarowości. Główne strategie redukcji wymiarowości. Rzutowanie. Uczenie rozmaitościowe. Analiza PCA. Zachowanie wariancji. Główne składowe. Techniki uczenia nienadzorowanego. Analiza skupień. Algorytm centroidów (Algorytm k-średnich). Granice algorytmu centroidów. Analiza skupień w segmentacji obrazu. Analiza skupień w przetwarzaniu wstępnym. Analiza skupień w uczeniu półnadzorowanym. Algorytm DBSCAN. Inne algorytmy analizy skupień. Mieszanki gaussowskie. Wykrywanie anomalii za pomocą mieszanin gaussowskich. Wyznaczanie liczby skupień. Modele bayesowskie mieszanin gaussowskich. Inne algorytmy służące do wykrywania anomalii i nowości. Ocena i poprawa jakości modeli. Reguła powrotu do średniej. Kryteria oceny modeli eksploracji danych. Ocena jakości modeli klasyfikacyjnych. Ocena jakości modeli regresyjnych. Ocena jakości modeli grupujących. Ocena jakości modeli rekomendujących. Ocena jakości modeli prognozujących. Poprawa jakości modeli. Znajdzenie optymalnej liczby klastrów. Wprowadzenie do sztucznych sieci neuronowych. Neurony biologiczne. Operacje logiczne przy użyciu neuronów. Perceptron. Perceptron wielowarstwowy i propagacja wsteczna. Regresyjne perceptrony wielowarstwowe. Klasyfikacyjne perceptrony wielowarstwowe. Głębokie uczenie. Dostrajanie hiperparametrów sieci neuronowej.	
Literatura podstawowa:	
- Géron A.: Uczenie maszynowe z użyciem Scikit-Learn i TensorFlow. Wydanie III, Helion, 2023 - Szeliga M.: Data science i uczenie maszynowe, PWN, 2017 - Morzy T.: Eksploracja danych – metody i algorytmy, PWN, 2024	
Literatura dodatkowa:	
Raschka S. i Mirjalili V.: Python. Uczenie maszynowe. Wydanie 2, Helion, 2019	

2. Moroney L.: Sztuczna inteligencja i uczenie maszynowe dla programistów. Praktyczny przewodnik po sztucznej inteligencji, Helion, 2021

Planowane formy/działania/metody dydaktyczne:

Wykład tradycyjny wspomagany jest technikami multimedialnymi. Ćwiczenia laboratoryjne – zajęcia praktyczne z wykorzystaniem wybranych narzędzi programowych. Na stronie internetowej prowadzącego zamieszczane są materiały z problemami i zadaniami laboratoryjnymi.

Sposoby weryfikacji efektów uczenia się osiągniętych przez studenta:

Efekty W_01 do W_03 zostaną zweryfikowane podczas egzaminu pisemnego w trakcie trwania sesji egzaminacyjnej. Na egzaminie pytania będą dotyczyły poznanych zagadnień. Przykładowe pytania egzaminacyjne:

- Przedstaw metodę liniowej klasyfikacji SVM.
- Przedstaw metodę klasyfikacji wieloetykietowej.
- Czym jest analiza skupień?

Przed egzaminem studenci będą mieli dostęp do przykładowych pytań.

Efekty U_01 do U_02 będą sprawdzane systematycznie na zajęciach laboratoryjnych w postaci wejściówek lub oceny wykonanej implementacji pod koniec zajęć.

Przykładowe zadania:

- Zaimplementuj model regresji liniowej dla podanych danych.
- Dokonaj klasyfikacji danych. Nadaj etykiety dla danych.
- Przeprowadź redukcję wielowymiarowości.

Forma i warunki zaliczenia:

Ocena z przedmiotu składa się z dwóch składowych:

- oceny z zajęć laboratoryjnych,
- oceny z egzaminu końcowego.

Na ocenę z zajęć laboratoryjnych składają się oceny cząstkowe uzyskane na regularnych zajęciach z nauczycielem akademickim, za które można uzyskać sumarycznie 40 pkt. Zaliczenie zajęć laboratoryjnych możliwe jest po uzyskaniu sumarycznie co najmniej 51% liczby punktów z tej formy zaliczenia.

W trakcie sesji odbędzie się egzamin końcowy. Do egzaminu mogą przystąpić osoby, które uzyskały zaliczenie laboratorium. Egzamin przewidziany jest w formie pisemnej. Można na nim uzyskać maksymalnie 60 pkt. Egzamin będzie zaliczony w przypadku uzyskania co najmniej 51% liczby punktów z tej formy zaliczenia. Ocena końcowa z przedmiotu, w zależności od sumy uzyskanych punktów (maksymalnie 100 pkt) jest następująca (w nawiasach ocena wg skali ECTS):

- 0 – 50 pkt: niedostateczna (F),
- 51 – 60 pkt: dostateczna (E),
- 61 – 70 pkt: dostateczna plus (D),
- 71 – 80 pkt: dobra (C),
- 81 – 90 pkt: dobra plus (B),
- 91 – 100 pkt: bardzo dobra (A).

Bilans punktów ECTS:

Studia stacjonarne

Aktywność	Obciążenie studenta
Udział w wykładach	21 godz.
Udział w ćwiczeniach laboratoryjnych	24 godz.
Przygotowanie się do egzaminu	25 godz.
Przygotowanie się do ćwiczeń laboratoryjnych	25 godz.
Udział w konsultacjach z przedmiotu	5 godz.
Sumaryczne obciążenie pracą studenta	100 godz.
Punkty ECTS za przedmiot	4 ECTS
Studia niestacjonarne	
Aktywność	Obciążenie studenta
Udział w wykładach	15 godz.
Udział w ćwiczeniach laboratoryjnych	15 godz.
Przygotowanie się do egzaminu	35 godz.
Przygotowanie się do ćwiczeń laboratoryjnych	30 godz.
Udział w konsultacjach z przedmiotu	5 godz.
Sumaryczne obciążenie pracą studenta	100 godz.
Punkty ECTS za przedmiot	4 ECTS