


KAPITAŁ LUDZKI
 NARODOWA STRATEGIA SPÓJNOŚCI

 Projekt współfinansowany przez
 Unię Europejską w ramach
 Europejskiego Funduszu
 Społecznego

UNIA EUROPEJSKA
 EUROPEJSKI
 FUNDUSZ SPOŁECZNY


Nazwa przedmiotu		Kod ECTS	
Uczenie maszynowe		11.0.0216	
Nazwa jednostki prowadzącej przedmiot			
Katedra Chemii i Radiochemii Środowiska			
Studia			
wydział	kierunek	poziom	pierwszego stopnia
Wydział Matematyki, Fizyki i Informatyki	Bioinformatyka	forma	stacjonarne
		moduł	Podstawowa
		specjalnościowy	Podstawowa
		specjalizacja	Podstawowa
Nazwisko osoby prowadzącej (osób prowadzących)			
prof. dr hab. Tomasz Puzyn; dr inż. Karolina Jagiełło; dr Alicja Mikołajczyk; dr Agnieszka Gajewicz-Skrętna			
Formy zajęć, sposób ich realizacji i przypisana im liczba godzin		Liczba punktów ECTS	
Formy zajęć		4	
Wykład, Ćw. laboratoryjne			
Sposób realizacji zajęć			
zajęcia w sali dydaktycznej			
Liczba godzin			
Ćw. laboratoryjne: 45 godz., Wykład: 15 godz.			
Termin realizacji przedmiotu			
2022/2023 zimowy			
Status przedmiotu		Język wykładowy	
obowiązkowy		polski	
Metody dydaktyczne		Forma i sposób zaliczenia oraz podstawowe kryteria oceny lub wymagania egzaminacyjne	
<ul style="list-style-type: none"> - •wykład - •ćwiczenia laboratoryjne: 		Sposób zaliczenia	
		<ul style="list-style-type: none"> - Zaliczenie na ocenę - Egzamin 	
		Formy zaliczenia	
		<ul style="list-style-type: none"> - egzamin ustny - egzamin pisemny z pytaniami (zadaniami) otwartymi - egzamin pisemny testowy - ustalenie oceny zaliczeniowej na podstawie ocen cząstkowych otrzymywanych w trakcie trwania semestru - •zaliczenie wykładu: egzamin pisemny (część testowa oraz zadania otwarte) oraz egzamin ustny (uzupełnienie egzaminu pisemnego) •zaliczenie ćwiczeń laboratoryjnych: ustalenie oceny zaliczeniowej na podstawie ocen cząstkowych (ocen uzyskanych z kolokwiiów i ocen za pisemne sprawozdania z wykonanych ćwiczeń) 	
		Podstawowe kryteria oceny	

- Wykład:
- Egzamin pisemny składający się z kilkunastu pytań testowych oraz kilku pytań otwartych (zadań) obejmujących zagadnienia wymienione w treściach programowych wykładu i ćwiczeń laboratoryjnych.
 - Warunkiem uzyskania pozytywnej oceny z zaliczenia pisemnego jest zdobycie minimum 51% punktów możliwych do uzyskania. Skala ocen jest zgodna z obowiązującym na Uniwersytecie Gdańskim regulaminem studiów.
 - Studenci, którzy uzyskali w pierwszym terminie zaliczenia pisemnego wynik 51% i więcej, a chcą podwyższyć ocenę, mogą zgłosić się na egzamin ustny. Ocena końcowa jest w tym przypadku średnią arytmetyczną z ocen uzyskanych na zaliczeniu pisemnym i ustnym.
 - Zaliczenie ustne jest obowiązkowe dla studentów, którzy uzyskali z egzaminu pisemnego wynik pomiędzy 41% a 50%. W tym przypadku student otrzymuje szansę uzupełnienia punktów brakujących do uzyskania oceny dostatecznej (omawia sposób poprawnego rozwiązania zadań z zaliczenia pisemnego). W tym przypadku nie ma możliwości poprawienia oceny z pierwszego terminu zaliczenia na wyższą.
 - Negatywna ocena z egzaminu (pisemnego i ustnego) musi być poprawiona podczas egzaminu poprawkowego odbywającego się w oparciu o te same zasady co egzamin w pierwszym terminie.
 - Ocena może być podwyższona o połowę studentom szczególnie aktywnie uczestniczącym w dyskusji naukowej podczas zajęć.
- Ćwiczenia laboratoryjne:
- Samodzielne wykonanie wszystkich zadanych ćwiczeń w pracowni komputerowej. Nieobecność można odrobić podczas zajęć z inną grupą ćwiczeniową lub w trakcie konsultacji u prowadzącego.
 - Potwierdzenie umiejętności prezentacji uzyskanych wyników oraz ich naukowej dyskusji poprzez uzyskanie pozytywnej oceny ze sprawozdań obejmujących wykonane ćwiczenia.
 - Zaliczenie wszystkich kolokwium wejściowych obejmujących podstawowe zagadnienia teoretyczne niezbędne do poprawnego wykonania ćwiczenia. Niezaliczone kolokwia należy poprawić w dodatkowym terminie wyznaczonym przez prowadzącego na zakończenie semestru (poza zajęciami).
 - Ocena końcowa z ćwiczeń jest średnią ważoną ze średnich arytmetycznych ocen otrzymanych z (i) kolokwium pisemnych (waga 40%), oraz (ii) sprawozdań obejmujących wykonane ćwiczenia (waga 60%).
 - Ocena może być podwyższona o połowę studentom szczególnie aktywnie uczestniczącym w dyskusji naukowej podczas zajęć.
 - Niezaliczenie ćwiczeń laboratoryjnych skutkuje niedopuszczeniem do zaliczenia wykładu do chwili uzyskania zaliczenia.

Sposób weryfikacji założonych efektów uczenia się

zakładany efekt kształcenia	konwersatorium	kolokwium	sprawozdanie	egzamin pisemny	egzamin ustny
	Wiedza				
KW_03	x		x	x	x
KW_04	x	x	x	x	x
	Umiejętności				
KU_02		x	x	x	x
KU_03		x	x	x	x
	Kompetencje				

Określenie przedmiotów wprowadzających wraz z wymogami wstępnymi

A. Wymagania formalne

1. Matematyka
2. Wstęp do informatyki
3. Programowanie
4. Techniki eksploracji danych wielowymiarowych

<p>B. Wymagania wstępne rachunek macierzowy, zagadnienie własne, badanie przebiegu zmienności funkcji, znajomość środowiska Linux, podstawy programowania przy wykorzystaniu języków obiektowych, metody uczenia maszynowego bez nadzoru</p>	
<p>Cele kształcenia</p> <ol style="list-style-type: none"> 1. Zaprezentowanie studentom zakresu możliwości zastosowania metod uczenia maszynowego w bioinformatyce. 2. Zdobycie przez studentów umiejętności posługiwania się najważniejszymi metodami uczenia maszynowego (poprawny dobór metody, wybór zmiennych, uczenie i walidacja modelu oraz interpretacja uzyskanych wyników). 3. Zapoznanie się przez studentów z dostępnym oprogramowaniem i językami skryptowymi realizującymi te metody, w szczególności: R, Python i KNIME. 	
<p>Treści programowe</p> <p>Modelowanie zależności z wykorzystaniem metod uczenia z nadzorem (regresyjnych i klasyfikacyjnych): regresja liniowa jednej i wielu zmiennych (LR i MLR), regresja logistyczna, regresja głównych składowych (PCR) oraz regresja metodą częściowych najmniejszych kwadratów (PLS); liniowa analiza dyskryminacyjna (LDA), nieliniowy klasyfikator k-najbliższych sąsiadów (kNN), naiwny klasyfikator bayesowski, drzewa decyzyjne (DT) i lasy losowe, drzewa klasyfikacji i regresji (CART), maszyny wektorów wspierających (SVM); sztuczne sieci neuronowe (ANN); Metody wyboru optymalnego zestawu zmiennych w modelu: wybór krokowy, algorytm eliminacji zmiennych nieinformatywnych (UVE), wybór przy użyciu algorytmu genetycznego (GA); walidacja modeli regresyjnych i klasyfikacyjnych (walidacja krzyżowa metodą „wyrzuc jeden” lub „wyrzuc kilka”, walidacja zewnętrzna). Wyznaczanie granic dziedziny zastosowania modeli.</p> <p>Na zajęciach szczególnie nacisk położony zostanie na praktyczne zastosowanie uczenia maszynowego do tworzenia modeli zależności w analizie danych genomicznych, transkryptomicznych, metabolomicznych, modeli ścieżek negatywnych skutków (ang. Adverse Outcome Pathways, AOP) oraz modeli typu struktura-aktywność biologiczna wykorzystywanych w farmacji i toksykologii. Zastosowanie technik uczenia maszynowego w medycynie spersonalizowanej. Tworzenie skryptów w Python. Łączenie dostępnych narzędzi (R, Python, KNIME).</p>	
<p>Wykaz literatury</p> <p>A. Literatura wymagana do ostatecznego zaliczenia zajęć (zdania egzaminu):</p> <p>A.1. wykorzystywana podczas zajęć:</p> <ul style="list-style-type: none"> • Skrypt do ćwiczeń laboratoryjnych przygotowywany przez pracowników Zespołu Chemometrii Środowiska <p>A.2. studiowana samodzielnie przez studenta</p> <ul style="list-style-type: none"> • J. Mazerski: Podstawy chemometrii. Gdańsk: Wydawnictwo Politechniki Gdańskiej, 2000 • M. Gałęwski: Programowanie w języku R. PWN, 2016 • M. Lutz: Python. Wprowadzenie. Helion, 2002 • S. Raschka: Python. Uczenie maszynowe. Helion, 2016 • https://www.youtube.com/user/KNIMETV • P. Biecek: Przewodnik po pakiecie R. Wrocław: Oficyna Wydawnicza GiS, 2014 • P. Biecek: Analiza danych z programem R. Modele liniowe z efektami stałymi, losowymi i mieszanymi: Warszawa PWN, 2020. <p>B. Literatura uzupełniająca</p> <ul style="list-style-type: none"> • S. D. Brown, R. Tauler, B. Walczak (red): Comprehensive chemometrics: Chemical and biochemical data analysis. Amsterdam: Elsevier, 2009 • R. Kramer: Chemometric techniques for quantitative analysis. New York: Marcel Dekker, Inc, 2005. 	
<p>Kierunkowe efekty uczenia się</p> <p>KW_03: Ma wiedzę z zakresu metod matematycznych i statystycznych pozwalającą na opis i modelowanie procesów i zjawisk biologicznych</p> <p>KW_04: Ma wiedzę w zakresie podstawowych technik i narzędzi badawczych stosowanych w bioinformatyce</p> <p>KU_02: Potrafi zastosować wiedzę z nauk przyrodniczych i ścisłych do formułowania, analizowania i rozwiązywania problemów związanych z bioinformatyką</p> <p>KU_03: Stosuje podstawowe metody matematyczne i statystyczne do opisu zjawisk i analizy danych; posiada umiejętność podstawowej analizy danych w profesjonalnych bazach danych wykorzystywanych w bioinformatyce</p>	<p>Wiedza</p> <ol style="list-style-type: none"> 1. Student zna podstawy teoretyczne (algorytm działania) najważniejszych metod uczenia maszynowego. 2. Student wskaże przykłady zastosowania metod uczenia maszynowego w bioinformatyce, farmacji i toksykologii. <p>Umiejętności</p> <ol style="list-style-type: none"> 1. Student potrafi poprawnie sformułować problem badawczy (pytanie badawcze) i dobrać do niego odpowiedni model teoretyczny oraz metodę uczenia maszynowego. 2. Student potrafi poprawnie zbudować model zależności i przeprowadzić jego walidację przy wykorzystaniu narzędzi dostępnych dla języka Python i samodzielnie zaprogramowane skrypty oraz poprawnie zinterpretować uzyskane wyniki. 3. Student potrafi poprawnie zaprezentować (w formie pisemnego sprawozdania) przedyskutować wyniki przeprowadzonego modelowania. <p>Kompetencje społeczne (postawy)</p> <p>Student określa zakres wiedzy i umiejętności w zakresie posługiwania się metodami uczenia maszynowego, w których chciałby się dalej rozwijać, biorąc pod uwagę plany zawodowe.</p>
<p>Kontakt</p>	

tomasz.puzyn@ug.edu.pl