Uniwersytet Gdański

Karta przedmiotu

Nazwa i kod przedmiotu	Uczenie maszynowe, PG_00193533		
Kierunek studiów	Bioinformatyka (O)		
Data rozpoczęcia studiów	październik 2026 r.	Rok akademicki realizacji przedmiotu	2028/2029
Poziom kształcenia	I stopnia - licencjackie	Grupa zajęć	Grupa zajęć obowiązkowych z zakresu kierunku studiów Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki
Forma studiów	stacjonarne	Sposób realizacji	na uczelni
Rok studiów	3	Język wykładowy	polski
Semestr studiów	5	Liczba punktów ECTS	4.0
Profil kształcenia	ogólnoakademicki	Forma zaliczenia	egzamin
Jednostka prowadząca	Rektor -> Wydział Chemii -> Katedra Chemii i Radiochemii Środowiska -> Pracownia Chemoinformatyki Środowiska		
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot	dr Artur Mirocki	
	Prowadzący zajęcia z przedmiotu		

Formy zajęć							
	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	45.0	0.0	0.0	60
	W tym liczba godzin zajęć na odległość: 0.0						

Aktywność studenta i liczba godzin pracy					
	Aktywność studenta	Udział w zajęciach dydaktycznych, objętych planem studiów	Udział w konsultacjach	Praca własna studenta	RAZEM
	Liczba godzin pracy studenta	60	0.0	40.0	100

Cel przedmiotu
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.

Efekty uczenia się przedmiotu			
	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[BIOINL3_W04] Ma zaawansowaną wiedzę w zakresie technik i narzędzi badawczych stosowanych w bioinformatyce	Student zna podstawy teoretyczne (algorytm działania) najważniejszych metod uczenia maszynowego. Student wskaże przykłady zastosowania metod uczenia maszynowego w bioinformatyce, farmacji i toksykologii.	[SW4] test/egzamin - ustny lub pisemny [SW1] wypowiedź ustna/rozmowa/ dyskusja [SW2] prezentacja/projekt/referat/ raport [SW3] opracowanie tekstowe/ praca pisemna
	[BIOINL3_W03] Ma wiedzę z zakresu metod matematycznych i statystycznych pozwalającą na opis i modelowanie procesów i zjawisk biologicznych	Student zna podstawy teoretyczne (algorytm działania) najważniejszych metod uczenia maszynowego. Student wskaże przykłady zastosowania metod uczenia maszynowego w bioinformatyce, farmacji i toksykologii.	[SW4] test/egzamin - ustny lub pisemny [SW1] wypowiedź ustna/rozmowa/ dyskusja [SW2] prezentacja/projekt/referat/ raport
	[BIOINL3_U03] Stosuje metody matematyczne i statystyczne do opisu zjawisk i analizy danych; posiada umiejętność analizy danych w profesjonalnych bazach danych wykorzystywanych w bioinformatyce	Student potrafi poprawnie sformułować problem badawczy (pytanie badawcze) i dobrać do niego odpowiedni model teoretyczny oraz metodę uczenia maszynowego. Student potrafi poprawnie zbudować model zależności i przeprowadzić jego walidację przy wykorzystaniu narzędzi dostępnych dla języka Phyton i samodzielnie zaprogramowane skrypty oraz poprawnie zinterpretować uzyskane wyniki. Student potrafi poprawnie zaprezentować (w formie pisemnego sprawozdania) przedyskutować wyniki przeprowadzonego modelowania.	[SU2] prezentacja/projekt/referat/ raport [SU3] opracowanie tekstowe/ praca pisemna [SU4] test/egzamin - ustny lub pisemny
	[BIOINL3_U02] Potrafi zastosować wiedzę z nauk przyrodniczych i ścisłych do formułowania, analizowania i rozwiązywania problemów związanych z bioinformatyką	Student potrafi poprawnie sformułować problem badawczy (pytanie badawcze) i dobrać do niego odpowiedni model teoretyczny oraz metodę uczenia maszynowego. Student potrafi poprawnie zbudować model zależności i przeprowadzić jego walidację przy wykorzystaniu narzędzi dostępnych dla języka Phyton i samodzielnie zaprogramowane skrypty oraz poprawnie zinterpretować uzyskane wyniki. Student potrafi poprawnie zaprezentować (w formie pisemnego sprawozdania) przedyskutować wyniki przeprowadzonego modelowania.	[SU2] prezentacja/projekt/referat/ raport [SU3] opracowanie tekstowe/ praca pisemna [SU4] test/egzamin - ustny lub pisemny

Treści przedmiotu	
	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 wyrzuć jeden lub wyrzuć kilka, walidacja zewnętrzna). Wyznaczanie granic dziedziny zastosowania modeli. Na zajęciach szczególny nacisk położony zostanie na praktyczne zastosowanie uczenia maszynowego do tworzenia modeli zależności w analizie danych genomicznych, transkryptomiczych, 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).

<table>
<tr><td>Wymagania wstępne i dodatkowe</td><td colspan="2">w.wstępne: rachunek macierzowy, zagadnienie własne, badanie przebiegu zmienności funkcji, znajomość środowiska Linux, podstawy programowania przy wykorzystaniu języków obiektowych, metody uczenia maszynowego bez nadzoru

w. formalne:

1. Matematyka
2. Wstęp do informatyki
3. Programowanie
4. Techniki eksploracji danych wielowymiarowych</td></tr>
<tr><td>Sposoby i kryteria oceniania osiąganych efektów uczenia się</td><td colspan="2"><table><tr><th>Sposób oceniania (składowe)</th><th>Próg zaliczeniowy</th><th>Składowa oceny końcowej</th></tr><tr><td>Egzamin</td><td>51.0%</td><td>100.0%</td></tr></table></td></tr>
<tr><td rowspan="3">Zalecana lista lektur</td><td>Podstawowa lista lektur</td><td>A. Literatura wymagana do ostatecznego zaliczenia zajęć (zdania egzaminu):

A.1. wykorzystywana podczas zajęć:

• Skrypt do ćwiczeń laboratoryjnych przygotowywany przez pracowników Zespołu Chemometrii Środowiska

A.2. studiowana samodzielnie przez studenta

• J. Mazerski: Podstawy chemometrii. Gdańsk: Wydawnictwo Politechniki Gdańskiej, 2000
• M. Gągolewski: 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.

B. Literatura uzupełniająca

• 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.</td></tr>
<tr><td>Uzupełniająca lista lektur</td><td>Literatura dodatkowa

• 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.</td></tr>
<tr><td>Adresy eZasobów</td><td></td></tr>
<tr><td>Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania</td><td colspan="2"></td></tr>
<tr><td>Praktyki zawodowe w ramach przedmiotu</td><td colspan="2">Nie dotyczy</td></tr>
</table>

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.