



Karta przedmiotu

Nazwa i kod przedmiotu	Metody eksploracji danych, PG_00064605						
Kierunek studiów	Fizyka Techniczna						
Data rozpoczęcia studiów	luty 2027 r.	Rok akademicki realizacji przedmiotu			2026/2027		
Poziom kształcenia	II stopnia	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	1	Język wykładowy			polski		
Semestr studiów	1	Liczba punktów ECTS			4.0		
Profil kształcenia	ogólnoakademicki	Forma zaliczenia			zaliczenie		
Jednostka prowadząca	Wydziały Politechniki Gdańskiej -> Wydział Fizyki Technicznej i Matematyki Stosowanej -> Katedra Fizyki Teoretycznej i Informatyki Kwant.						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr inż. Michał Piłat				
	Prowadzący zajęcia z przedmiotu						
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	0.0	0.0	30.0	45
	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	45		5.0		50.0	100
Cel przedmiotu	Celem przedmiotu jest przedstawienie studentom zaawansowanych metod służących do analizy dużych zbiorów danych oraz wykształcenie krytycznego podejścia do otrzymanych wyników. Studenci nauczą się również przygotowywania prezentacji zawierających treści teoretyczne i odpowiednie przykłady.						

Efekty uczenia się przedmiotu	Efekt kierunkowy	Efekt z przedmiotu	Sposób weryfikacji i oceny efektu
	[K7_W04] posiada pogłębioną znajomość metod matematycznych, numerycznych i symulacyjnych stosowanych przy opisie i modelowaniu zjawisk fizycznych.	Posiada pogłębioną znajomość metod matematycznych, numerycznych i symulacyjnych stosowanych przy opisie i modelowaniu danych	[SW2] Ocena wiedzy zawartej w prezentacji
	[K7_U06] potrafi identyfikować i oceniać zagrożenia, efektywność ekonomiczną oraz przydatność proponowanych rozwiązań inżynierskich. Posiada umiejętność krytycznej oceny zaproponowanych lub istniejących rozwiązań biorąc pod uwagę także czynniki pozatechniczne, w tym aspekty etyczne.	Potrafi zastosować zdobytą wiedzę z zakresu fizyki do zagadnień z obszaru innych nauk ścisłych, nauk przyrodniczych lub technicznych.	[SU1] Ocena realizacji zadania
	[K7_U01] potrafi uczyć się samodzielnie przez całe życie, pozyskiwać i integrować informacje z literatury, baz danych oraz innych właściwie dobranych źródeł. Posiada umiejętność krytycznej analizy i selekcji informacji, korzysta z zasobów informacji patentowej.	Potrafi uzyskać informację na temat technik wykorzystywanych w w eksploracji danych, dowiadywać się, jakie nowe modele są stosowane i ocenić na ile dany model może być zastosowany do danego przypadku.	[SU2] Ocena umiejętności analizy informacji
	[K7_U09] potrafi w środowisku akademickim i poza nim komunikować się na tematy z zakresu fizyki i nauk pokrewnych, w tym informatyki stosowanej, fizyki stosowanej i fotowoltaiki, a także organizować i prowadzić merytoryczną debatę oraz ukierunkowywać innych na nieustanne poszukiwanie rzetelnej wiedzy.	Potrafi poprowadzić krótką prezentację na temat wybranego zagadnienia z metod eksploracji danych	[SU5] Ocena umiejętności zaprezentowania wyników realizacji zadania

Treści przedmiotu	<p>Treści przedmiotu - wykład</p> <ol style="list-style-type: none"> 1. Definicja, etapy i cele eksploracji danych. Wybrane struktury i modele używane w eksploracji danych. Działanie algorytmów i ich stosowanie w dużych zbiorach danych. 2. Wybrane pojęcia statystyki opisowej. W tym testy i przedziały ufności 3. Zrozumienie i czyszczenie danych. Eliminacja zbędnych zmiennych. Problem współliniowości. Metody wyboru zmiennych objaśniających i ich zastosowanie. Eksploracyjna analiza danych 4. Modelowanie regresji liniowej. Metoda najmniejszych kwadratów. Współczynniki determinacji i korelacji. 5. Modelowanie regresji wielokrotnej. Tworzenie modelu i przeprowadzanie wnioskowania. 6. Regresja logistyczna. Oszacowanie największej wiarygodności. Interpretacja wyników i wnioskowanie co do wiarygodności zmiennych objaśniających. 7. Estymacja bayesowska i sieci bayesowskie. Klasyfikacja maksymalnego <i>a posteriori</i>. Iloraz szans <i>a posteriori</i>. Równoważenie danych. Naiwny klasyfikator bayesowski 8. Metody nadzorowane i nienadzorowane. Metodologia modelowania nadzorowanego. Algorytm k-najbliższych sąsiadów 9. Drzewa decyzyjne. Algorytmy C4.5 i CART w zastosowaniu do rzeczywistych danych. Reguły decyzyjne. 10. Sieci neuronowe. Kodowanie sygnałów wejściowych i wyjściowych. Sigmoidalna funkcja aktywacji. Współczynnik uczenia. 11. Sieci samoorganizujące się. Sieci Kohonena. Zastosowanie do zadania grupowania 12. Reguły asocjacyjne. Analiza podobieństw i koszyka sklepowego. Metoda uogólnionej redukcji reguł. J-miara 13. Metody grupowania hierarchicznego. Algorytm k-średnich 14. Ograniczona maszyna Boltzmanna. Struktura. Trenowanie. 15. Techniki ewaluacji modeli do zadań opisu, szacowania, przewidywania i klasyfikacji. Współczynnik błędu. Falszywe klasyfikacje.
	<p>Treści przedmiotu - seminarium</p> <ol style="list-style-type: none"> 1. Sieci Kohonena 2. Reguły asocjacyjne 3. Techniki ewaluacji modelu 4. Metody redukcji wymiaru zagadnienia 5. Regresja logistyczna 6. Algorytmy genetyczne 7. Modele regresji 8. Modelowanie z wykorzystaniem regresji wielorakiej 9. Metody eksploracji danych w marketingu 10. Odkrywanie asocjacji 11. Odkrywanie wzorów sekwencji 12. Grupowanie 13. Dekompozycja SVD 14. Uczenie nienadzorowane: tworzenie klastrów 15. Uczenie nadzorowane: metody statystyczne 16. Uczenie nadzorowane: drzewa decyzyjne, algorytmy reguł i ich hybrydy

	<p>17. Eksploracja tekstów</p> <p>18. Ocena modeli danych</p> <p>19. Bezpieczeństwo, prywatność i wyszukiwanie danych</p> <p>20. Podstawy matematyczne: algebra liniowa</p> <p>21. Podstawy matematyczne: prawdopodobieństwo</p> <p>22. Podstawy matematyczne: linie i płaszczyzny</p> <p>23. Podstawy matematyczne: zbiory</p> <p>24. Podstawy matematyczne: wybrane testy statystyczne, np. test t-Studenta, test f (Fishera)</p> <p>25. Metody identyfikacji obserwacji odstających</p> <p>26. Analiza kanoniczna</p> <p>27. Ekstrakcja cech</p> <p>28. Klasyfikacja, np. algorytm Las losowy (Random forest)</p>									
Wymagania wstępne i dodatkowe	Podstawy matematyki, w tym opisowej statystyki. Podstawy z fizyki. Dociekliwość i krytycyzm.									
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	<table border="1"> <thead> <tr> <th>Sposób oceniania (składowe)</th> <th>Próg zaliczeniowy</th> <th>Składowa oceny końcowej</th> </tr> </thead> <tbody> <tr> <td>Ocena realizacji zadania</td> <td>50.0%</td> <td>50.0%</td> </tr> <tr> <td>Ocena prezentacji</td> <td>50.0%</td> <td>50.0%</td> </tr> </tbody> </table>	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej	Ocena realizacji zadania	50.0%	50.0%	Ocena prezentacji	50.0%	50.0%
Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej								
Ocena realizacji zadania	50.0%	50.0%								
Ocena prezentacji	50.0%	50.0%								
Zalecana lista lektur	<table border="1"> <tr> <td>Podstawowa lista lektur</td> <td> <p>Literatura:</p> <p>Daniel T. Larose, Metody i modele eksploracji danych, PWN</p> <p>Tadeusz Morzy, Eksploracja danych. Metody i algorytmy, PWN</p> </td> </tr> <tr> <td>Uzupełniająca lista lektur</td> <td>Pozycje internetowe.</td> </tr> <tr> <td>Adresy eZasobów</td> <td></td> </tr> </table>	Podstawowa lista lektur	<p>Literatura:</p> <p>Daniel T. Larose, Metody i modele eksploracji danych, PWN</p> <p>Tadeusz Morzy, Eksploracja danych. Metody i algorytmy, PWN</p>	Uzupełniająca lista lektur	Pozycje internetowe.	Adresy eZasobów				
Podstawowa lista lektur	<p>Literatura:</p> <p>Daniel T. Larose, Metody i modele eksploracji danych, PWN</p> <p>Tadeusz Morzy, Eksploracja danych. Metody i algorytmy, PWN</p>									
Uzupełniająca lista lektur	Pozycje internetowe.									
Adresy eZasobów										
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	<p>1. Opisz podobieństwa i różnice pomiędzy sieciami neuronowymi, sieciami Kohonena i ograniczonymi maszynami Boltzmana. 2.. Jakich porad udzieliłabyś/udzieliłbyś osobie przystępującej do eksploracji danych? 3. Zbuduj drzewo decyzyjne za pomocą dowolnie wybranego algorytmu wyznaczającego ryzyko kredytowe na podstawie podanych danych:</p>									
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy									