



Karta przedmiotu

Nazwa i kod przedmiotu	Metody eksploracji danych, PG_00045761						
Kierunek studiów	Fizyka Techniczna						
Data rozpoczęcia studiów	luty 2023 r.		Rok akademicki realizacji przedmiotu		2022/2023		
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	Wydział Fizyki Technicznej i Matematyki Stosowanej -> Katedra Fizyki Teoretycznej i Informatyki Kwantowej						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		prof. dr hab. Józef Sienkiewicz				
	Prowadzący zajęcia z przedmiotu		prof. dr hab. Józef Sienkiewicz dr hab. inż. Maciej Demianowicz				
Formy zajęć i metody nauczania	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	30.0	0.0	0.0	0.0	30.0	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	5.0	35.0	100		
Cel przedmiotu	<b>Uczyć studentów podstawowych i zaawansowanych metod służących do analizy dużych zbiorów danych. Nauczyć krytycznego podejścia do otrzymywanych wyników. Uczyć przygotowywania prezentacji zawierających treści teoretyczne i odpowiednie przykłady.</b>						
Efekty uczenia się przedmiotu	Efekt kierunkowy		Efekt z przedmiotu		Sposób weryfikacji i oceny efektu		
	[K7_U06] Potrafi zastosować zdobytą wiedzę z zakresu fizyki do zagadnień z obszaru innych nauk ścisłych, nauk przyrodniczych lub technicznych.		Potrafi zastosować zdobytą wiedzę z zakresu fizyki do zagadnień z obszaru innych nauk ścisłych, nauk przyrodniczych lub technicznych.		[SU1] Ocena realizacji zadania		
	[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		

Treści przedmiotu	<p><b>Wykład:</b></p> <ol style="list-style-type: none"> <li>1. Definicja, etapy i cele eksploracji danych. Podstawowe struktury i modele używane w eksploracji danych. Działanie algorytmów i ich stosowanie w dużych zbiorach danych.</li> <li>2. Podstawowe pojęcia statystyki opisowej. W tym testy i przedziały ufności</li> <li>3. Zrozumienie i czyszczenie danych. Eliminacja zbędnych zmiennych. Problem współliniowości. Metody wyboru zmiennych objaśniających i ich zastosowanie. Eksploracyjna analiza danych</li> <li>4. Modelowanie regresji liniowej. Metoda najmniejszych kwadratów. Współczynniki determinacji i korelacji.</li> <li>5. Modelowanie regresji wielokrotnej. Tworzenie modelu i przeprowadzanie wnioskowania.</li> <li>6. Regresja logistyczna. Oszacowanie największej wiarygodności. Interpretacja wyników i wnioskowanie co do wiarygodności zmiennych objaśniających.</li> <li>7. Estymacja bayesowska i sieci bayesowskie. Klasyfikacja maksymalnego <i>a posteriori</i>. Iloraz szans <i>a posteriori</i>. Równoważenie danych. Naiwny klasyfikator bayesowski</li> <li>8. Metody nadzorowane i nienadzorowane. Metodologia modelowania nadzorowanego. Algorytm k-najbliższych sąsiadów</li> <li>9. Drzewa decyzyjne. Algorytmy C4.5 i CART w zastosowaniu do rzeczywistych danych. Reguły decyzyjne.</li> <li>10. Sieci neuronowe. Kodowanie sygnałów wejściowych i wyjściowych. Sigmoidalna funkcja aktywacji. Współczynnik uczenia.</li> <li>11. Sieci samoorganizujące się. Sieci Kohonena. Zastosowanie do zadania grupowania</li> <li>12. Reguły asocjacyjne. Analiza podobieństw i koszyka sklepowego. Metoda uogólnionej redukcji reguł. J-miara</li> <li>13. Metody grupowania hierarchicznego. Algorytm k-średnich</li> <li>14. Ograniczona maszyna Boltzmana. Struktura. Trenowanie.</li> <li>15. Techniki ewaluacji modeli do zadań opisu, szacowania, przewidywania i klasyfikacji. Współczynnik błędu. Falszywe klasyfikacje.</li> </ol> <p><b>Seminarium:</b></p> <p>Referaty z wybranych problemów eksploracji danych.</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 prezentacji</td> <td>50.0%</td> <td>50.0%</td> </tr> <tr> <td>Ocena realizacji zadania</td> <td>50.0%</td> <td>50.0%</td> </tr> </tbody> </table>	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej	Ocena prezentacji	50.0%	50.0%	Ocena realizacji zadania	50.0%	50.0%
Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej								
Ocena prezentacji	50.0%	50.0%								
Ocena realizacji zadania	50.0%	50.0%								

Zalecana lista lektur	Podstawowa lista lektur	<b>Literatura:</b>  Daniel T. Larose, <i>Metody i modele eksploracji danych</i> , PWN  Tadeusz Morzy, <i>Eksploracja danych. Metody i algorytmy</i> , PWN
	Uzupełniająca lista lektur	<b>Pozycje internetowe.</b>
	Adresy eZasobów	Adresy na platformie eNauczanie:
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	<b>1. Opisz podobieństwa i różnice pomiędzy sieciami neuronowymi, sieciami Kohonena i ograniczonymi maszynami Boltzmana.</b> <b>2.. Jakich porad udzieliłabyś/udzieliłbyś osobie przystępującej do eksploracji danych?</b> <b>3. Zbuduj drzewo decyzyjne za pomocą dowolnie wybranego algorytmu wyznaczającego ryzyko kredytowe na podstawie podanych danych:</b>	
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy	