Uniwersytet Gdański

Karta przedmiotu

| | | | |
|---|---|---|---|
| Nazwa i kod przedmiotu | Podstawy uczenia maszynowego, PG_00198500 | | |
| Kierunek studiów | Informatyka (O) | | |
| Data rozpoczęcia studiów | październik 2026 r. | Rok akademicki realizacji przedmiotu | 2027/2028 |
| Poziom kształcenia | I stopnia - licencjackie | Grupa zajęć | Grupa zajęć obowiązkowych z zakresu kierunku studiów<br>Grupa zajęć powiązanych z prowadzonymi badaniami naukowymi w dziedzinie nauki związanej z kierunkiem - profil ogólnoakademicki |
| Forma studiów | stacjonarne | Sposób realizacji | na uczelni |
| Rok studiów | 2 | Język wykładowy | polski |
| Semestr studiów | 4 | Liczba punktów ECTS | 5.0 |
| Profil kształcenia | ogólnoakademicki | Forma zaliczenia | zaliczenie |
| Jednostka prowadząca | Rektor -> Wydział Matematyki, Fizyki i Informatyki -> Instytut Informatyki -> Zakład Informatyki Kwantowej | | |
| Imię i nazwisko wykładowcy (wykładowców) | Odpowiedzialny za przedmiot | dr inż. Paweł Mazurek | |
| | Prowadzący zajęcia z przedmiotu | | |

| Formy zajęć | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Forma zajęć | Wykład | Ćwiczenia | Laboratorium | Projekt | Seminarium | RAZEM |
| | Liczba godzin zajęć | 30.0 | 0.0 | 30.0 | 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 | 65.0 | 125 |

| Cel przedmiotu | |
|---|---|
| | 1. Zapoznanie z podstawowymi metodami uczenia maszynowego<br><br>2. Nabycie umiejętności stosowania podstawowych algorytmów uczenia maszynowego - praca w środowisku Python<br><br>3. Opanowanie matematycznych metod modelowania problemów i narzędzi uczenia maszynowego oraz oceny ich przydatności |

**Efekty uczenia się przedmiotu**

| Efekt kierunkowy | Efekt z przedmiotu | Sposób weryfikacji i oceny efektu |
|---|---|---|
| [INFOL3_W06] zna i rozumie zaawansowane modele systemów baz danych, ze szczególnym uwzględnieniem modelu relacyjnego | tudent zna i rozumie rolę relacyjnych baz danych w procesie uczenia maszynowego. Potrafi pozyskiwać dane z relacyjnych baz danych, rozumie ich strukturę oraz wykorzystuje dane zapisane w modelu relacyjnym do przygotowania zbiorów uczących, ich przetwarzania i analizy. | [SW4] test/egzamin - ustny lub pisemny<br>[SW2] prezentacja/projekt/referat/ raport |
| [INFOL3_U02] potrafi wykorzystywać posiadaną wiedzę z zakresu matematyki wyższej w modelowaniu i rozwiązywaniu złożonych problemów | Student potrafi wykorzystywać wiedzę z zakresu matematyki wyższej, w szczególności algebry liniowej, rachunku różniczkowego, rachunku prawdopodobieństwa i statystyki do modelowania danych, analizy i implementacji algorytmów uczenia maszynowego oraz rozwiązywania złożonych problemów predykcyjnych i klasyfikacyjnych. | [SU2] prezentacja/projekt/referat/ raport<br>[SU4] test/egzamin - ustny lub pisemny |
| [INFOL3_U06] potrafi dobrać i stosować właściwe metody i narzędzia informatyczne do rozwiązywania złożonych problemów, potrafi wykorzystać posiadaną wiedzę poprzez właściwy dobór źródeł i informacji z nich pochodzących, dokonywanie oceny, krytycznej analizy i syntezy tych informacji | Student potrafi dobrać odpowiednie metody uczenia maszynowego, takie jak regresja, klasyfikacja, klasteryzacja, drzewa decyzyjne, SVM czy sieci neuronowe, oraz stosuje właściwe narzędzia informatyczne, np. Python, biblioteki scikit-learn, pandas, NumPy lub środowiska typu Jupyter Notebook. | [SU2] prezentacja/projekt/referat/ raport<br>[SU4] test/egzamin - ustny lub pisemny |
| [INFOL3_W03] zna i rozumie zaawansowane zagadnienia z zakresu algorytmów i struktur danych, języków formalnych, teorii automatów i złożoności obliczeniowej oraz sztucznej inteligencji | 1. Student/ka potrafi wykorzystać środowisko Python do uczenia maszynowego.<br>2. Student/ka zna metody analizy danych.<br>3. Student/ka potrafi wykorzystać narzędzia biblioteki TensorFlow do uczenia sieci neuronowych. | [SW4] test/egzamin - ustny lub pisemny<br>[SW2] prezentacja/projekt/referat/ raport |

**Treści przedmiotu**

1.Wprowadzenie do podstawowej terminologii i notacji. Strategia tworzenia systemu w uczeniu maszynowym 2. Python w uczeniu maszynowym .Wykorzystywanie środowiska Python do uczenia maszynowego: Anaconda, Jupyter, NumPy, Pandas, Matplotlib, SciPy.3.Wstępne przetwarzanie danych. Najlepsze metody oceny modelu i strojenie parametryczne. 4. Analiza danych za pomocą analizy regresyjnej, analiza skupień 5.Uczenie sieci neuronowych za pomocą biblioteki TensorFlow 6. Przykłady zastosowania : (i) Modelowanie danych sekwencyjnych za pomocą rekurencyjnych sieci neuronowych (ii) Praca z tekstem (iii) Praca z obrazami

**Wymagania wstępne i dodatkowe**

Znajomość podstaw programowania, algorytmów i złożoności, języków i paradygmatów programowania obiek
językach.

**Sposoby i kryteria oceniania osiąganych efektów uczenia się**

| Sposób oceniania (składowe) | Próg zaliczeniowy | Składowa oceny końcowej |
|---|---|---|
| projekt | 51.0% | 70.0% |
| egzamin testowy z treści wykładowych | 51.0% | 30.0% |

**Zalecana lista lektur**

Podstawowa lista lektur


1. Python Machine Learning By Example, Yuxi Liu, Packt Publishing, 2024

2. M. Szeliga, Data science i uczenie maszynowe, Wydawnictwo naukowe PWN, Warszawa, 2017.

3. T. Morzy, Eksploracja danych metody i algorytmy, Wydawnictwo naukowe PWN, Warszawa, 2013.

4. Christopher M. Bishop Pattern Recognition and Machine Learning, Springer, 2007

| | | |
|---|---|---|
| | Uzupełniająca lista lektur | 1. J. Koronacki, J. Ćwik: Statystyczne systemy uczące się. Wydanie drugie, EXIT, Warszawa, 20072. K. Krawiec, J.Stefanowski, Uczenie maszynowe i sieci neuronowe. Wydawnictwo Politechniki Poznańskiej, Poznań, 2004,3. P. Cichosz, Systemy uczace sie, WNT, Warszawa, 2000.4. W. Duch, J. Korbicz, L.Rutkowski,R. Tadeusiewicz, Sieci neuronowe, Exit, Warszawa, 20005. M. Gągolewski, M. Bartoszuk, A. Cena, Przetwarzanie i analiza danych w języku Python, PWN, 20166. W. McKinney, Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython, O'Reilly Media, 20127. W. Richert, L.P. Coelho, Building Machine Learning Systems with Python, Packt Publishing, 20138. M. Lutz, Learning Python, O'Reilly Media, 20139.. E. Bressert, SciPy and NumPy, O'Reilly Media, 2012 |
| | Adresy eZasobów | |
| Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania | 1. Typy uczenia maszynowego (z nazdorem, bez nadzoru, ze wzmocnieniem).<br>2. Definicje przeuczenia i niedouczenia, wykrywanie.<br>3. Walidacja krzyżowa.<br>4. Regularyzacja.<br>5. Selekcja cech i redukcja wymiarowości.<br>6. Inżynieria danych, w tym one-hot-encoding i gęste kodowanie.<br>7. Miary poprawności predykcji (precyzja, recall, AUC, ...)<br>8. Teoria Bayesa i zastosowanie w uczeniu maszynowym, najlepsze praktyki z nią związane (np. jak sobie radzić, gdy nierównowaga pomiędzy klasami; wygładzanie; co zrobić, gdy zbiory danych są małe/duże), typowe parametry metod<br>9. Algorytmy konstrukcji drzew decyzyjnych. Algorytmy konstrukcji lasów. Algorytm boostingu drzew. Typowe parametry drzew, lasów i boostowanych drzew. Wady i zalety drzew, lasów i boostowanych drzew.<br>10. Miary informacji (współczynnik Giniego, entropia,...)<br>11. Regresji logistyczna jako problem optymalizacyjny. Istnienie rozwiązań analitycznych w zależności od zastosowanej regularyzacji. Typowe parametry.<br>12. Regresja logistyczna w zastosowaniu do klasyfikacji wieloklasowej.<br>13. Regresja liniowa i drzewa regresyjne. Różnice względem drzew i regresji klasyfikacyjnej.<br>14. Sieci neuronowe, typowe parametry. Funkcje aktywacji. Metody przeciwdziałania przeuczeniu (dropout, wczesne zatrzymanie). Najlepsze praktyki strojenia sieci.<br>15. Tokenizacja, tagowanie, stemming, lematyzacja. Stopwords.<br>16. Wizualizacja obiektów metodą t-SNE.<br>17. Redukcja wymiarowości metodą PCA.<br>18. Bag-of-words, tf-idf vs. gęste kodowanie - zalety i wady, typowe zastosowania.<br>19. Algorytm k-średnich. Typowe parametry, najlepsze praktyki.<br>20. Zaawansowane metody ustalania tematu tekstu (NMF, LDA).<br>21. Metoda wektorów nośnych jako problem optymalizacji półokreślonej.<br>22. Zastosowanie maszyny wektorów nośnych w klasyfikacji wieloklasowej.<br>23. Jądra liniowe i nieliniowe w maszynie wektorów nośnych. Zastosowania i najlepsze praktyki. Typowe parametry.<br>24. Konwolucyjne sieci neuronowe: struktura, typowe parametry, porównanie z niekonwolucyjnymi sieciami neuronowymi. Źródła nieliniowości i konwolucyjności, funkcje grupujące. Najlepsze praktyki. | |
| Praktyki zawodowe w ramach przedmiotu | Nie dotyczy | |

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.