

Karta przedmiotu

Nazwa i kod przedmiotu	MACHINE LEARNING, PG_00070557						
Kierunek studiów	Analityka gospodarcza						
Data rozpoczęcia studiów	październik 2026 r.		Rok akademicki realizacji przedmiotu		2027/2028		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć specjalnościowych 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	3		Liczba punktów ECTS		3.0		
Profil kształcenia	ogólnoakademicki		Forma zaliczenia		zaliczenie		
Jednostka prowadząca	Wydziały Politechniki Gdańskiej -> Wydział Zarządzania i Ekonomii -> Katedra Statystyki i Ekonometrii						
Imię i nazwisko wykładowcy (wykładowców)	Odpowiedzialny za przedmiot		dr hab. Michał Pietrzak				
	Prowadzący zajęcia z przedmiotu						
Formy zajęć	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	15.0	0.0	30.0	0.0	0.0	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		3.0		27.0	75
Cel przedmiotu	zapoznanie studentów z podstawowymi koncepcjami, metodami i algorytmami uczenia maszynowego, umożliwiającymi tworzenie modeli analizy danych, prognozowania i podejmowania decyzji w oparciu o dane. Studenci zdobędą wiedzę teoretyczną oraz praktyczne umiejętności w zakresie stosowania wybranych algorytmów nadzorowanych i nienadzorowanych, przetwarzania danych, walidacji modeli oraz oceny ich skuteczności. Kurs kładzie nacisk na zrozumienie procesu budowy modeli ML, ich interpretację i zastosowanie w rzeczywistych problemach z różnych dziedzin.						
Efekty uczenia się przedmiotu	Efekt kierunkowy		Efekt z przedmiotu		Sposób weryfikacji i oceny efektu		
	[K7_U01] potrafi tworzyć innowacyjne rozwiązania złożonych i nieustrukturyzowanych procesów uwzględniając nieprzewidywalne warunki otoczenia poprzez syntezę informacji pochodzących z różnych źródeł.		potrafi projektować i wdrażać modele uczenia maszynowego do analizy złożonych procesów przy zmiennych warunkach otoczenia. Potrafi ocenić przydatność różnych źródeł danych i integrować je w celu uzyskania spójnych i trafnych wniosków.		[SU2] Ocena umiejętności analizy informacji [SU4] Ocena umiejętności korzystania z metod i narzędzi [SU1] Ocena realizacji zadania		
	[K7_W03] demonstuje pogłębioną wiedzę w zakresie zastosowań metod analitycznych oraz technik formułowania i rozwiązywania problemów społeczno-ekonomicznych.		posiada zaawansowaną wiedzę na temat metod analizy danych i algorytmów uczenia maszynowego, umożliwiającą identyfikację oraz modelowanie złożonych i nieustrukturyzowanych procesów. Zna współczesne narzędzia i technologie wykorzystywane do przetwarzania danych pochodzących z heterogenicznych źródeł.		[SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym		
	[K7_K01] jest gotów do krytycznej oceny swojej wiedzy z zakresu analityki gospodarczej oraz poszukiwania opinii ekspertów w przypadku trudności z samodzielnym rozwiązaniem problemu.		jest gotów do krytycznej oceny modeli i analiz uczenia maszynowego, rozpoznając ograniczenia stosowanych metod oraz w sytuacjach złożonych sięgając po wsparcie eksperckie.		[SK3] Ocena umiejętności organizacji pracy [SK5] Ocena umiejętności rozwiązywania problemów występujących w praktyce		

Treści przedmiotu	Treści przedmiotu - wykład		
	1. Wprowadzenie do uczenia maszynowego definicje, zakres i znaczenie ML, klasyfikacja metod (uczenie nadzorowane, nienadzorowane, ze wzmocnieniem) oraz przegląd środowiska obliczeniowego Python i bibliotek ML 2. Podstawy matematyczne dla uczenia maszynowego elementy algebry liniowej, rachunku różniczkowego i całkowego, statystyki matematycznej oraz metod optymalizacji wykorzystywanych w modelowaniu 3. Przygotowanie danych metody czyszczenia danych, transformacje i przekształcenia cech, eksploracyjna analiza danych (EDA) oraz techniki wizualizacji 4. Modele regresyjne regresja liniowa i jej rozszerzenia, regresja logistyczna, interpretacja parametrów modeli oraz metody regularyzacji (Ridge, Lasso) 5. Metody klasyfikacji przegląd podstawowych algorytmów klasyfikacyjnych oraz ich zastosowań w problemach decyzyjnych 6. Walidacja modeli metody oceny jakości modeli, podział danych, walidacja krzyżowa oraz analiza zjawisk overfittingu i underfittingu 7. Uczenie nienadzorowane metody klasteryzacji, redukcji wymiarowości oraz wykrywania anomalii w danych 8. Strojenie modeli metody doboru hiperparametrów oraz koncepcja pipeline'ów w procesie uczenia modeli 9. Uczenie głębokie podstawy teoretyczne i architektura sieci neuronowych 10. Zastosowania uczenia maszynowego przykłady zastosowań oraz etapy realizacji projektu ML (pipeline) 11. Etyka sztucznej inteligencji zagrożenia biasu, interpretowalności modeli oraz odpowiedzialności w systemach AI		
	Treści przedmiotu - laboratoria		
	1. Konfiguracja środowiska obliczeniowego instalacja i konfiguracja środowiska Python, zarządzanie zależnościami (venv/conda) oraz wykorzystanie środowiska Jupyter Notebook 2. Biblioteki NumPy i pandas operacje na strukturach danych, przetwarzanie i manipulacja danymi tabelarycznymi 3. Eksploracyjna analiza danych wczytywanie danych, analiza struktury zbiorów danych oraz identyfikacja braków i typów zmiennych 4. Wstępne przetwarzanie danych (preprocessing) czyszczenie danych, imputacja braków, kodowanie zmiennych kategorycznych oraz skalowanie cech 5. Wykrywanie anomalii metody identyfikacji obserwacji odstających 6. Wizualizacja danych zastosowanie narzędzi graficznych do analizy rozkładów i zależności między zmiennymi 7. Regresja liniowa implementacja modelu, estymacja parametrów oraz interpretacja wyników 8. Rozszerzone modele regresyjne regresja wielomianowa oraz metody regularyzacji (Ridge, Lasso) 9. Regresja logistyczna budowa i zastosowanie modelu klasyfikacyjnego dla zmiennych binarnych 10. Algorytmy KNN i Naive Bayes implementacja oraz analiza porównawcza metod klasyfikacyjnych 11. Drzewa decyzyjne i Random Forest modele drzewiaste oraz metody zespołowe 12. Maszyny wektorów nośnych (SVM) klasyfikacja z wykorzystaniem funkcji jądra i analiza granic decyzyjnych 13. Walidacja modeli podział danych, walidacja krzyżowa oraz ocena stabilności modeli 14. Zjawisko overfittingu analiza wpływu złożoności modelu na zdolność generalizacji 15. Klasteryzacja K-means implementacja i analiza grupowania danych 16. Metody klasteryzacji DBSCAN i hierarchiczne algorytmy gęstościowe i aglomeracyjne 17. Analiza głównych składowych (PCA) redukcja wymiarowości i analiza wariancji 18. Strojenie hiperparametrów zastosowanie metod Grid Search i Random Search		
Wymagania wstępne i dodatkowe	Statystyka opisowa, statystyka matematyczna, ekonometria, prognozowanie, podstawy programowania w R/Pythonie.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Projekt końcowy	60.0%	50.0%
	Kolokwium w laboratorium komputerowym	60.0%	50.0%

Zalecana lista lektur	Podstawowa lista lektur	1. Rogalski, M. (2019). <i>Uczenie maszynowe z Pythonem</i>. Gliwice: Helion. 2. Raschka, S., & Mirjalili, V. (2017). <i>Python machine learning</i>. Birmingham, UK: Packt Publishing. 3. Chollet, F. (2021). <i>Deep learning with Python</i> (2nd ed.). Shelter Island, NY: Manning Publications. 4. Géron, A. (2022). <i>Hands-on machine learning with scikit-learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems</i> (3rd ed.). Sebastopol, CA: O'Reilly Media. 5. Bishop, C. M. (2006). <i>Pattern recognition and machine learning</i>. New York, NY: Springer. 6. Murphy, K. P. (2012). <i>Machine learning: A probabilistic perspective</i>. Cambridge, MA: MIT Press. 7. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). <i>An introduction to statistical learning: With applications in R</i> (2nd ed.). New York, NY: Springer. 8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). <i>Deep learning</i>. Cambridge, MA: MIT Press. 9. Burkov, A. (2019). <i>The hundred-page machine learning book</i>. Quebec City, Canada: Andriy Burkov. 10. Hastie, T., Tibshirani, R., & Friedman, J. (2009). <i>The elements of statistical learning: Data mining, inference, and prediction</i> (2nd ed.). New York, NY: Springer.
	Uzupełniająca lista lektur	1. Kotu, V., & Deshpande, B. (2019). <i>Data science: Concepts and practice</i> (2nd ed.). Cambridge, MA: Morgan Kaufmann. 2. McKinney, W. (2022). <i>Python for data analysis: Data wrangling with pandas, NumPy, and IPython</i> (3rd ed.). Sebastopol, CA: O'Reilly Media. 3. VanderPlas, J. (2016). <i>Python data science handbook: Essential tools for working with data</i>. Sebastopol, CA: O'Reilly Media. 4. Molnar, C. (2022). <i>Interpretable machine learning</i> (2nd ed.). N.p.: Lulu.com. 5. Provost, F., & Fawcett, T. (2013). <i>Data science for business: What you need to know about data mining and data-analytic thinking</i>. Sebastopol, CA: O'Reilly Media. 6. Grus, J. (2019). <i>Data science from scratch: First principles with Python</i> (2nd ed.). Sebastopol, CA: O'Reilly Media. 7. Brownlee, J. (2020). <i>Machine learning mastery with Python</i>. Melbourne, Australia: Machine Learning Mastery. 8. Sarker, I. H. (2021). <i>Machine learning: Algorithms and applications</i>. Boca Raton, FL: CRC Press. 9. Taddy, M. (2019). <i>Business data science: Combining machine learning and economics to optimize, automate, and accelerate business decisions</i>. New York, NY: McGraw-Hill.
	Adresy eZasobów	
Przykładowe zagadnienia/ przykładowe pytania/ realizowane zadania	1. Skonfiguruj środowisko Python (venv/conda), zainstaluj wymagane biblioteki oraz uruchom Jupyter Notebook, aby przygotować środowisko do analizy danych. 2. Wczytaj rzeczywisty zbiór danych (np. CSV), przeprowadź jego eksploracyjną analizę (EDA) oraz zidentyfikuj brakujące wartości i potencjalne anomalie w danych. 3. Przygotuj dane do modelowania poprzez imputację braków, zakodowanie zmiennych kategorycznych oraz skalowanie cech, aby uzyskać zbiór gotowy do uczenia maszynowego. 4. Zbuduj model regresji liniowej do przewidywania zmiennej ciągłej, a następnie oceń jego jakość przy użyciu metryk (MSE, R^2) i zinterpretuj wyniki. 5. Rozszerz model regresji o cechy wielomianowe oraz zastosuj regularyzację (Ridge lub Lasso), aby poprawić zdolność generalizacji modelu. 6. Zaimplementuj modele klasyfikacyjne (regresja logistyczna, KNN, Naive Bayes) i porównaj ich skuteczność na tym samym zbiorze danych. 7. Zbuduj modele Decision Tree, Random Forest oraz SVM, a następnie przeanalizuj wpływ ich parametrów na jakość predykcji. 8. Zastosuj podział danych oraz walidację krzyżową, aby ocenić stabilność modelu i zidentyfikować problem overfittingu. 9. Przeprowadź klasteryzację danych metodą K-means oraz DBSCAN, a następnie porównaj uzyskane grupowania i ich interpretację. 10. Zastosuj analizę PCA w celu redukcji wymiarowości danych, a następnie oceń wpływ redukcji na wyniki klasteryzacji lub klasyfikacji.	
Zajęcia praktyczne w ramach przedmiotu	Nie dotyczy	