

Karta przedmiotu

Nazwa i kod przedmiotu	MACHINE LEARNING, PG_00068305						
Kierunek studiów	Analityka gospodarcza						
Data rozpoczęcia studiów	październik 2025 r.		Rok akademicki realizacji przedmiotu		2026/2027		
Poziom kształcenia	II stopnia		Grupa zajęć		Grupa zajęć fakultatywnych 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		4.0		
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 inż. Karol Flisikowski				
	Prowadzący zajęcia z przedmiotu						
Formy zajęć i metody nauczania	Forma zajęć	Wykład	Ćwiczenia	Laboratorium	Projekt	Seminarium	RAZEM
	Liczba godzin zajęć	0.0	0.0	30.0	0.0	0.0	30
	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	30		5.0		65.0	100
Cel przedmiotu	Celem przedmiotu jest 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_W03] demonstruje pogłębioną wiedzę w zakresie zastosowań metod analitycznych oraz technik formułowania i rozwiązywania problemów społeczno-ekonomicznych.		Ma 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_U01] potrafi tworzyć innowacyjne rozwiązania złożonych i nieustrukturyzowanych procesów uwzględniając nieprzewidywalne warunki otoczenia poprzez syntezę informacji pochodzących z różnych źródeł.		Student 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.			[SU1] Ocena realizacji zadania [SU4] Ocena umiejętności korzystania z metod i narzędzi [SU2] Ocena umiejętności analizy informacji	

1.
Wprowadzenie do uczenia maszynowego
 - Definicja i historia uczenia maszynowego
 - Różnica między sztuczną inteligencją, uczeniem maszynowym i głębokim
 - Obszary zastosowań uczenia maszynowego (rozpoznawanie obrazów, analiza tekstu, prognozy itp.)
 - Podstawowe kategorie algorytmów ML: nadzorowane, nienadzorowane, ze wzmocnieniem
2.
Podstawy matematyczne uczenia maszynowego
 - Wprowadzenie do algebry liniowej (macierze, wektory, operacje)
 - Statystyka: miary centralne, rozkłady prawdopodobieństwa, estymacja
 - Pojęcia związane z optymalizacją (funkcje kosztu, gradient descent)
3.
Przygotowanie danych
 - Podstawy przetwarzania danych: czyszczenie, uzupełnianie braków, normalizacja, standaryzacja
 - Przekształcanie zmiennych: kodowanie zmiennych kategorycznych, inżynieria cech
 - Podział danych: zbiór treningowy, walidacyjny, testowy
 - Wyzwania związane z dużymi zbiorami danych (Big Data)
4.
Uczenie nadzorowane (Supervised Learning)
 - **Regresja:** liniowa, wielomianowa, regresja logistyczna
 - **Klasyfikacja:** Naive Bayes, k-NN, drzewa decyzyjne, SVM
 - **Modele ensemble:** Random Forest, Gradient Boosting (XGBoost, LightGBM, CatBoost)
 - **Neural Networks (MLP)**
 - Optymalizacja: Cross-validation, regularizacja (L1, L2), walidacja krzyżowa
5.
Uczenie nienadzorowane (Unsupervised Learning)
 - **Klasteryzacja:** K-Means, DBSCAN, aglomeracyjna klasteryzacja hierarchiczna
 - **Redukcja wymiarowości:** PCA, t-SNE, LDA
 - **Analiza zależności:** analiza głównych składowych, analiza czynnikowa

6.	Uczenie głębokie (Deep Learning) • Wprowadzenie do sieci neuronowych • Budowa i działanie perceptronów • Sieci konwolucyjne (CNN) do analizy obrazów • Sieci rekurencyjne (RNN) i LSTM w analizie sekwencyjnej • Transfer learning i fine-tuning • Wprowadzenie do bibliotek: TensorFlow, Keras, PyTorch
7.	Optymalizacja modeli i hyperparameter tuning • Wybór najlepszych hiperparametrów modelu (Grid Search, Random Search, Bayesian Optimization) • Regularizacja i techniki redukcji nadmiernego dopasowania (dropout, early stopping) • Cross-validation i walidacja krzyżowa • Analiza błędów i ocena wyników (RMSE, MAE, AUC-ROC, F1-score)
8.	Zastosowania uczenia maszynowego w praktyce • Rozpoznawanie obrazów: klasyfikacja obrazów z wykorzystaniem CNN • Przetwarzanie języka naturalnego (NLP): analiza tekstów, modele word embeddings (Word2Vec, GloVe) • Rekomendacje: algorytmy filtracji oparte na treści i współpracy (Collaborative Filtering) • Predykcja: prognozowanie szeregów czasowych i analiza trendów • Anomalie i detekcja oszustw: identyfikacja nieprawidłowości w danych
9.	Etyka i odpowiedzialność w uczeniu maszynowym • Problemy związane z uprzedzeniami w danych (bias) • Etyczne wykorzystanie algorytmów w różnych branżach • Transparentność i interpretowalność modeli ML • Odpowiedzialność za decyzje oparte na algorytmach (np. w medycynie, finansach, wymiarze sprawiedliwości)

	10. Warsztaty praktyczne i projekty - Implementacja wybranych algorytmów ML w Pythonie (scikit-learn, pandas, numpy, matplotlib) - Analiza rzeczywistych przypadków (case studies) - Praca z danymi z różnych źródeł (np. dane zebrane z mediów społecznościowych, dane finansowe, obrazy) - Mini-projekt: wdrożenie i prezentacja wyników na podstawie własnego zbioru danych		
Wymagania wstępne i dodatkowe	Statystyka opisowa, statystyka matematyczna, podstawy programowania w R/Pythonie.		
Sposoby i kryteria oceniania osiągniętych efektów uczenia się	Sposób oceniania (składowe)	Próg zaliczeniowy	Składowa oceny końcowej
	Wejściówki	50.0%	10.0%
	Sprawozdania	50.0%	40.0%
	Projekt końcowy	50.0%	50.0%
Zalecana lista lektur	Podstawowa lista lektur	1. Rogalski, M. (2019). <i>Uczenie maszynowe z Pythonem.</i> Helion. 2. Raschka, S., & Mirjalili, V. (2020). <i>Python. Machine Learning i Deep Learning. Biblioteki scikit-learn i TensorFlow 2. Wydanie III.</i> Helion. 3. Siemieniuk, S. (2018). <i>Python w uczeniu maszynowym. Praktyczne podejście.</i> Wydawnictwo Helion.	
	Uzupełniająca lista lektur	1. Chollet, F. (2025). <i>Deep learning with Python</i> (3rd ed.). Manning Publications. 2. Chen, S., Zhang, H., & Li, J. (2024). <i>Deep learning and machine learning Python data structures and mathematics fundamentals: From theory to practice.</i> Springer.	
	Adresy eZasobów	Podstawowe https://github.com/ageron/handson-ml3 - Seria notebooków Jupyter, które wprowadzą Cię w podstawy uczenia maszynowego i głębokiego w Pythonie z wykorzystaniem Scikit-Learn, Keras i TensorFlow 2. https://github.com/ageron/handson-ml3 - Python Machine Learning (3rd Ed.) Code Repository Uzupełniające Adresy na platformie eNauczanie:	

Zagadnienia:

1.
Podstawy uczenia maszynowego:
 - Klasyfikacja vs. regresja: czym się różnią i kiedy je stosować?
 - Przypadki użycia różnych algorytmów klasyfikacji (np. drzewa decyzyjne, SVM, KNN).
 - Podstawowe techniki optymalizacji w uczeniu maszynowym (np. gradient descent).
 2.
Przygotowanie danych:
 - Przeszukiwanie i oczyszczanie danych (missing values, outliers).
 - Normalizacja, standaryzacja i inżynieria cech.
 - Metody selekcji cech (feature selection) i redukcji wymiarów (PCA, LDA).
 3.
Uczenie głębokie:
 - Podstawy neuronów i sieci neuronowych.
 - Architektury sieci neuronowych: CNN, RNN, GANs.
 - Overfitting i regularizacja w głębokim uczeniu (dropout, L2 regularization).
 4.
Modelowanie i ocena:
 - Cross-validation i techniki walidacji.
 - Matryce pomyłek, Precision, Recall, F1-score.
 - Strojenie hiperparametrów i wybór najlepszego modelu.
- Pytania:**
1. Jakie są różnice między klasycznymi algorytmami uczenia maszynowego, a metodami głębokiego uczenia?
 2. Jakie techniki można zastosować w celu zapobiegania overfittingowi w modelach głębokiego uczenia?
 3. W jaki sposób można ocenić skuteczność modelu klasyfikacyjnego w kontekście nierównych klas?
 4. Co to jest gradient descent i jak wpływa na proces uczenia modeli?

Realizowane zadania:

	1. Zadanie 1: - Wykorzystanie Scikit-learn do klasyfikacji danych z wykorzystaniem algorytmu KNN. Porównanie wyników z innymi algorytmami, np. SVM, drzewa decyzyjne. 2. Zadanie 2: - Zastosowanie algorytmu PCA do redukcji wymiarów w zbiorze danych (np. Iris dataset) i ocena efektywności tej metody na wynikach klasyfikacji. 3. Zadanie 3: - Budowa modelu sieci neuronowej przy użyciu Keras/TensorFlow do klasyfikacji obrazów (np. dataset CIFAR-10). Porównanie wyników z klasycznymi algorytmami ML. 4. Zadanie 4: - Przeprowadzenie analizy i oczyszczania danych (missing values, outliers) w zbiorze danych o klientach banku. Następnie budowa modelu predykcji bankructwa klienta. 5. Zadanie 5: - Zastosowanie metody regularizacji (L2, Dropout) w modelu głębokiego uczenia do rozwiązania problemu klasyfikacji obrazów. Ocena skuteczności w zapobieganiu overfittingowi.
Praktyki zawodowe w ramach przedmiotu	Nie dotyczy

Dokument wygenerowany elektronicznie. Nie wymaga pieczęci ani podpisu.