

**KARTA KURSU (realizowanego w module specjalności)**  
**Data Science**

Nazwa	<b>Zaawansowane metody uczenia maszynowego</b>
Nazwa w j. ang.	Advanced Methods of Machine Learning

Koordynator	Dr hab Jozef Kapusta, prof. UKEN	Zespół dydaktyczny
		Dr hab Jozef Kapusta, prof. UKEN
Punktacja ECTS*	3	

Opis kursu (cele kształcenia)

Celem kursu jest pogłębienie wiedzy o zaawansowanych algorytmach ML, zastosowanie nowoczesnych metod w analizie danych i rozwiązywaniu problemów praktycznych oraz rozwijanie umiejętności implementacji, optymalizacji i oceny modeli ML. Kurs przygotowuje studentów do samodzielnej pracy nad projektami analitycznymi z wykorzystaniem rzeczywistych zbiorów danych.

Warunki wstępne

Wiedza	Wymagana znajomość: podstawy uczenia maszynowego i statystyki, znajomość języka Python i bibliotek ML (np. scikit-learn, TensorFlow/PyTorch), podstawy algebry liniowej i analizy danych.
Umiejętności	Wymagane umiejętności: implementacja podstawowych algorytmów ML, analiza i wizualizacja danych, interpretacja wyników modeli i ocena ich jakości, samodzielne rozwiązywanie problemów analitycznych.
Kursy	

Efekty uczenia się

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Wiedza	<p>Po zakończeniu kursu student:</p> <p>W01: zna istotę i podstawy matematyczne stosowane w uczeniu maszynowym, w tym optymalizacji, statystyki, rachunku prawdopodobieństwa i algebry liniowej, niezbędne do analizy i interpretacji danych;</p> <p>W02: posiada pogłębioną wiedzę na temat kluczowych metod uczenia maszynowego, w tym klasyfikacji danych o standardowej i złożonej strukturze oraz głębokiego uczenia, i rozumie ich praktyczne zastosowania;</p> <p>W03: zna narzędzia i metody eksploracji danych oraz komunikacji wyników, potrafi analizować i wizualizować dane oraz wyciągać wnioski zgodnie z zasadami etyki i ochrony danych.</p>	<p>SD_W01 SD_W03 - SD_W06 SD_W09</p>

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Umiejętności	Po zakończeniu kursu student:	
	<p>U01: potrafi zaprojektować i zaimplementować modele uczenia maszynowego i głębokich sieci neuronowych, dobrać odpowiednie metody i algorytmy w zależności od typu danych oraz przeprowadzić analizę ich skuteczności i ograniczeń;</p> <p>U02: potrafi dokonać eksploracyjnej analizy danych rzeczywistych, w tym wizualizacji i interpretacji wyników, a także zaprezentować rozwiązanie w sposób czytelny dla zespołu, uwzględniając aspekty etyczne i ochrony danych.</p>	<p>SD_U02 SD_U03 SD_U05 SD_U06 SD_U08 SD_U09</p>

	Efekt uczenia się dla kursu	Odniesienie do efektów kierunkowych
Kompetencje społeczne	Po zakończeniu kursu student:	
	<p>K01: ma świadomość społecznej i zawodowej roli specjalisty zajmującego się analizą danych i uczeniem maszynowym oraz znaczenia wpływu swojej pracy na otoczenie;</p> <p>K02: potrafi formułować i uzasadniać opinie dotyczące metod eksploracji danych, uczenia maszynowego, sztucznej inteligencji i technologii chmurowych, uwzględniając aspekty praktyczne i teoretyczne;</p> <p>K03: ma świadomość znaczenia profesjonalnego i etycznego zachowania w pracy z danymi, potrafi identyfikować i rozstrzygać dylematy etyczne związane z projektami z obszaru Data Science.</p>	SD_K01 - SD_K03

#### Studia stacjonarne

Organizacja							
Forma zajęć	Wykład (W)	Ćwiczenia w grupach					
		A	K	L	S	P	E
Liczba godzin	15			30			

#### Studia niestacjonarne

Organizacja							
Forma zajęć	Wykład (W)	Ćwiczenia w grupach					
		A	K	L	S	P	E
Liczba godzin	10			20			

## Opis metod prowadzenia zajęć

Kurs prowadzony jest w formie zajęć teoretycznych i praktycznych, łącząc wykłady, laboratoria z zadaniami indywidualnymi oraz projekty grupowe.

## Formy sprawdzania efektów uczenia się

	E – learning	Gry dydaktyczne	Ćwiczenia w szkole	Zajęcia terenowe	Praca laboratoryjna	Projekt indywidualny	Projekt grupowy	Udział w dyskusji	Referat	Praca pisemna (esej)	Egzamin ustny	Egzamin pisemny	Inne
W01					X	X	X	X					
W02					X	X	X	X					
W03					X	X	X	X					
U01					X	X	X	X					
U02					X	X	X	X					
K01								X					
K02								X					
K03								X					

Kryteria oceny	<p>Ocena końcowa zależy od: ocen cząstkowych z laboratoriów i testów, systematyczności realizacji zadań, jakości projektu zespołowego lub indywidualnego.</p> <p>W szczególności:</p> <p>Ocena dobra lub bardzo dobra z laboratoriów wymaga:</p> <ul style="list-style-type: none"> <li>• samodzielnej implementacji algorytmów ML,</li> <li>• analizy warunków stosowalności i interpretacji wyników,</li> <li>• umiejętności walidacji i strojenia modeli.</li> </ul> <p>Ocena dobra lub bardzo dobra z projektu wymaga:</p> <ul style="list-style-type: none"> <li>• samodzielnej analizy danych z użyciem zaawansowanych metod,</li> <li>• interpretacji wyników i wyciągania wniosków,</li> <li>• przygotowania czytelnego raportu i prezentacji.</li> </ul> <p>Dodatkowo oceniana jest aktywność i systematyczność na zajęciach.</p>
----------------	--

Uwagi	
-------	--

## Treści merytoryczne (wykaz tematów)

1. Wprowadzenie i zaawansowane zagadnienia ML. Przegląd klasycznych algorytmów i ich ograniczeń. Wyzwania w dużych zbiorach danych i wysokiej wymiarowości. Trendy i nowoczesne kierunki w uczeniu maszynowym (deep learning, ensemble methods, probabilistic ML).
2. Drzewa decyzyjne i metody zespołowe. Algorytm budowy drzew, pruning, overfitting. Random Forest - teoria agregacji i bagging. Boosting - AdaBoost, Gradient Boosting, XGBoost, LightGBM. Interpretacja modeli i analiza znaczenia cech.
3. Metody kernelowe i Support Vector Machines. Podstawy SVM - hiperplan, margines, wsparcie wektorów. Kernel trick i funkcje jądra (RBF, polynomial). Zastosowanie w klasyfikacji, regresji i redukcji wymiarów.
4. Modele probabilistyczne i bayesowskie. Naiwny klasyfikator Bayesa, założenia i ograniczenia. Gaussian Processes - regresja probabilistyczna. Bayesian optimization - teoria i zastosowania w strojeniu modeli.
5. Głębokie uczenie i sieci neuronowe. Architektury sieci: perceptron wielowarstwowy, CNN, RNN, LSTM, Transformers. Funkcje aktywacji, propagacja wsteczna, optymalizacja. Regularizacja, dropout, batch normalization. Transfer learning - teoria i zastosowanie.

6. Uczenie nienadzorowane i reprezentacyjne. Redukcja wymiarów: PCA, t-SNE, UMAP. Klasteryzacja: k-means, hierarchiczna, DBSCAN. Autoenkodery – teoria i interpretacja reprezentacji ukrytej.
7. Optymalizacja i walidacja modeli. Walidacja krzyżowa, metryki ewaluacji dla klasyfikacji i regresji. Optymalizacja hiperparametrów: grid search, random search, Bayesian optimization. Analiza błędów i diagnostyka modeli
8. Interpretowalność i etyka w ML. Metody wyjaśnialności: SHAP, LIME. Analiza znaczenia cech i wpływu danych wejściowych. Fairness, bias, odpowiedzialność i etyczne aspekty wdrażania modeli ML.

#### Wykaz literatury podstawowej

1. Aurélien Géron – „Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow”, 3rd Edition, O'Reilly, 2022
2. Christopher M. Bishop – „Pattern Recognition and Machine Learning”, Springer, 2006
3. Ian Goodfellow, Yoshua Bengio, Aaron Courville – „Deep Learning”, MIT Press, 2016

#### Wykaz literatury uzupełniającej

1. Trevor Hastie, Robert Tibshirani, Jerome Friedman – „The Elements of Statistical Learning”, 2nd Edition, Springer, 2009
2. Sebastian Raschka, Vahid Mirjalili – „Python Machine Learning”, 3rd Edition, Packt, 2019

#### Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) **studia stacjonarne**

liczba godzin w kontakcie z prowadzącymi	Wykład	15
	Konwersatorium (ćwiczenia, laboratorium itd.)	30
	Pozostałe godziny kontaktu studenta z prowadzącym	10
liczba godzin pracy studenta bez kontaktu z prowadzącymi	Lektura w ramach przygotowania do zajęć	10
	Przygotowanie krótkiej pracy pisemnej lub referatu po zapoznaniu się z niezbędną literaturą przedmiotu	
	Przygotowanie projektu lub prezentacji na podany temat (praca w grupie)	10
	Przygotowanie do egzaminu/zaliczenia	
Ogółem bilans czasu pracy		75
Liczba punktów ECTS w zależności od przyjętego przelicznika		3

#### Bilans godzinowy zgodny z CNPS (Całkowity Nakład Pracy Studenta) **studia niestacjonarne**

liczba godzin w kontakcie z prowadzącymi	Wykład	10
	Konwersatorium (ćwiczenia, laboratorium itd.)	20
	Pozostałe godziny kontaktu studenta z prowadzącym	15
liczba godzin pracy studenta bez kontaktu z prowadzącymi	Lektura w ramach przygotowania do zajęć	15
	Przygotowanie krótkiej pracy pisemnej lub referatu po zapoznaniu się z niezbędną literaturą przedmiotu	
	Przygotowanie projektu lub prezentacji na podany temat (praca w grupie)	15
	Przygotowanie do egzaminu/zaliczenia	
Ogółem bilans czasu pracy		75
Liczba punktów ECTS w zależności od przyjętego przelicznika		3