



# Centrum Studiów Zaawansowanych Politechniki Warszawskiej

Pl. Politechniki 1, 00-661 Warszawa, tel./fax +48 22 234 6003 (6002), www.csz.pw.edu.pl



Uczelniana Oferta Studiów Zaawansowanych SYLABUS 2018/2019	
Nazwa przedmiotu (jęz. polski i angielski)	Przetwarzanie i analiza danych w języku Python (PADP) Python for Data Processing and Analysis
Liczba punktów ECTS	3 ECTS

Osoby prowadzące	Tytuł naukowy	Imię i nazwisko	Katedra / Instytut/ Centrum/ Inne
	dr hab. inż.	Marek Gągolewski	MiNI PW, IBS PAN
Osoba odpowiedzialna za przedmiot	dr hab. inż.	Marek Gągolewski	MiNI PW, IBS PAN

Semestr studiów	<i>Zimowy 2018</i>
Typ przedmiotu (możliwości wyboru) obowiązkowy O fakultatywny F	F
Wymagania wstępne	Podstawy programowania strukturalnego Podstawy rachunku prawdopodobieństwa Podstawy statystyki matematycznej
Poziom przedmiotu Podstawowy P Średniozaawansowany Ś Zaawansowany Z	Ś
Charakter zajęć, liczba godzin w semestrze, liczba godzin w tygodniu. 1) podać rodzaj prowadzonych zajęć dla danego przedmiotu: wykłady (W); ćwiczenia (Ć); laboratorium (L); projekt (P) 2) podać liczbę godzin w tygodniu np. W - 2; Ć - 2; L - 3; P - 0 3) podać liczbę godzin w semestrze np. W - 30; Ć - 30; L - 45; P - 0	<i>W – 2 godz. w tygodniu, łącznie: 30 godzin</i>
Sugerowana liczba godzin pracy własnej	45 godzin obejmuje : 30 godzin przygotowywanie się słuchacza do wykładów, 15 – przygotowywanie się

	stuchacza do egzaminu.
<b>Całkowita liczba godzin:</b>	75 godzin
<b>Aspekty międzynarodowe (jeśli są)</b>	<i>dokumentacja techniczna wyłącznie w języku angielskim</i>
<b>Język wykładowy</b>	polski
<b>Cel przedmiotu</b>	Uczestnicy kursu mają możliwość dogłębnego poznania technik programowania w języku Python oraz najbardziej przydatnych pakietów z punktu widzenia przetwarzania, wizualizacji i analizy danych (np. podczas opracowywania wyników własnych badań empirycznych lub symulacyjnych do publikacji). Nabywają nie tylko umiejętność stosowania, ale także samodzielnej implementacji wybranych algorytmów uczenia maszynowego (np. sieci neuronowych) m.in. przy użyciu wysokopoziomowych operacji na tensorach. W trakcie wykładu zostaną omówione najpopularniejsze algorytmy analizy skupień, klasyfikacji i regresji.
<b>Treść przedmiotu</b>	
<ol style="list-style-type: none"> <li>1. Wprowadzenie do języka Python 3 i środowiska Jupyter/IPython</li> <li>2. Podstawy programowania w języku Python. Typy skalarne.</li> <li>3. Typy sekwencyjne i iterowalne, słowniki, zbiory</li> <li>4. Instrukcje sterujące, funkcje</li> <li>5. Podstawowe polecenia w powłoce (bash). Skrypty, moduły, pakiety</li> <li>6. Programowanie obiektowe</li> <li>7. Obliczenia na wektorach, macierzach i innych tensorach (NumPy oraz Theano lub TensorFlow, także na GPU)</li> <li>8. Ramki danych i najważniejsze operacje na nich (Pandas)</li> <li>9. Wizualizacja danych (matplotlib, Seaborn)</li> <li>10. Przegląd metod wnioskowania statystycznego (SciPy, statsmodels)</li> <li>11. Przegląd algorytmów uczenia maszynowego w zadaniach regresji, klasyfikacji i analizy skupień (scikit-learn)</li> </ol>	
<b>Spis zalecanych lektur</b>	
<b>LP.</b>	<b>Autor, Tytuł, Wydawnictwo,</b>
1.	Gagolewski M., Bartoszek M., Cena A., Przetwarzanie i analiza danych w języku Python, Wydawnictwo Naukowe PWN, Warszawa, 2016
2.	McKinney W., Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython, O'Reilly Media, 2012
3.	Hastie T., Tibshirani R., Friedman J., The elements of Statistical Learning, Springer, 2017
4.	Bishop C.M., Pattern Recognition and Machine Learning, Springer, 2006
5.	Koronacki J., Ćwik J., Statystyczne systemy uczące się, EXIT, 2008

<b>Metody oceny</b> ( ocena, egz. pisemny, egz. ustny, projekt)	<b>projekt</b>
--	----------------

<b>Uwagi dodatkowe</b>	Zajęcia odbędą się, jeżeli zapisze się co najmniej 20 osób. Jako jedyną formą zaliczenia przedmiotu są oceny.
------------------------	---

**Tabela 1. Efekty kształcenia**

Numer (symbol)	Efekty kształcenia słuchacza, który zaliczył przedmiot, potrafi	Sposób weryfikacji osiągnięcia efektu
<b>WIEDZA</b>		
<b>PADP_W1</b>	Zna składnię języka Python 3 oraz wysokopoziomowe operacje na wektorach, macierzach i innych tensorach oraz ramkach danych	Projekt
<b>PADP_W2</b>	Zna podstawowe klasy, metody i funkcje udostępniane przez pakiety NumPy, SciPy, scikit-learn, Pandas, matplotlib, seaborn, scikit-learn, statsmodels	Projekt
<b>PADP_W3</b>	Ma uporządkowaną wiedzę w zakresie podstawowych metod uczenia maszynowego	Projekt
<b>UMIEJĘTNOŚCI</b>		
<b>PADP_U1</b>	Umie wykorzystać dokumentację techniczną bibliotek i innych narzędzi programistycznych w języku angielskim do implementacji programów.	Projekt
<b>PADP_U2</b>	<b>Umie samodzielnie zaimplementować algorytmy analizy danych w języku Python.</b>	Projekt
<b>PADP_U3</b>	<b>Umie wykorzystać gotowe algorytmy analizy danych dostępne w pakietach języka Python.</b>	Projekt
<b>PADP_U4</b>	<b>Umie stosować techniki przygotowywania zbiorów danych do ich analizy.</b>	Projekt
<b>KOMPETENCJE</b>		
<b>PADP_K1</b>	Rozumie konieczność dalszego samokształcenia	Obserwacja na zajęciach, egzamin
<b>PADP_K2</b>	Rozumie znaczenie metod interdyscyplinarnych w nauce	Obserwacja na zajęciach.