

## TREŚCI KSZTAŁCENIA

Nazwa studiów podyplomowych: „**Data science w praktyce**”

Wymiar kształcenia (sem.): dwa semestry

### CHARAKTERYSTYKA TREŚCI KSZTAŁCENIA

#### **1. Wprowadzenie do języka Python**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest nabycie umiejętności uruchomienia środowiska Python zarówno w systemach Windows jak i inux jak również zdobycie wiedzy na temat architektury oprogramowania pisanego z wykorzystaniem języka Python oraz umiejętność praktycznego zastosowania tej wiedzy w projekcie.

Treści merytoryczne:

- instalacja środowiska Python w systemie Windows i Linux,
- instalacja i importowanie bibliotek Python,
- wprowadzenie do środowiska PyCharm,
- podstawowe elementy języka Python: ciągi tekstowe, listy, tablice jedno i wielowymiarowe, instrukcje warunkowe, pętle, obsługa plików,
- podział kodu programu na moduły, klasy i funkcje. Składnia wg wytycznych PEP8,
- projekt.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe elementy składni języka Python; sposób efektywnego wykorzystania zewnętrznych pakietów;

*umiejętności:* (potrafi) zainstalować i przygotować środowisko pod wytyczne projektu; napisać prosty program z wykorzystaniem języka Python w środowisku PyCharm;

*kompetencje społeczne:* (jest gotów do) rozumienia znaczenia ciągłego dokształcania się; rozumienia znaczenia współpracy z innymi członkami zespołu przy realizacji projektu programistycznego.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG7, SP\_P7S\_WK2, SP\_P7S\_UW4, SP\_P7S\_UK1, SP\_P7S\_UK3,  
SP\_P7S\_UU1,

SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 3

#### **2. Warsztat badacza danych**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest nabycie umiejętności obsługi systemu Linux zarówno poprzez GUI oraz wiersz poleceń jak również dobytec umiejętności zarządzania repozytoriami plików za pomocą systemu Git.

Treści merytoryczne:

- instalacja systemu Linux w środowisku wirtualnym,
- podstawowa obsługa systemu Linux poprzez GUI,
- podstawowe komendy Bash,

- uprawnienia oraz zarządzanie użytkownikami i grupami,
- instalacja oprogramowania i pakietów,
- język znaczników Markdown,
- wykorzystanie systemu kontroli wersji Git,
- Środowisko IPython notebook.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe polecenia wiersza poleceń systemu Linux; zasady wykorzystywania systemu Git do zarządzania dowolnym zbiorem plików;

*umiejętności:* (potrafi) zainstalować i skonfigurować system Linux; stworzyć repozytorium plików i współdzielić je z innymi użytkownikami;

*kompetencje społeczne:* (jest gotów do) rozumienia znaczenia ciągłego dokształcania się; rozumienia znaczenia współdzielenia swoich efektów pracy z innymi w celu pogłębiania wiedzy i podnoszenia kompetencji.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG7, SP\_P7S\_UW1, SP\_P7S\_UW4, SP\_P7S\_UK1, SP\_P7S\_UK2, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1, SP\_P7S\_KO1, SP\_P7S\_KO3

*Liczba ECTS:* 1

### **3. Wizualizacja i eksploracja danych**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest zapoznanie słuchacza z technikami przeprowadzania eksploracji danych oraz zaznajomienie słuchacza z metodami graficznej prezentacji danych.

*Treści merytoryczne:*

- biblioteka numpy,
- przetwarzanie danych - biblioteka pandas,
- normalizacja i przetwarzanie danych - zmiana formatu, missing values, itp,
- generowanie wykresów – matplotlib, pandas.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) rozmaite techniki prezentacji danych; zalety i ograniczenia poznanych technik prezentacji; gotowe biblioteki do analizy danych;

*umiejętności:* (potrafi) dobrać odpowiednią technikę prezentacji do otrzymanych wyników; przygotować zestawienie danych w postaci graficznej, a następnie dokonać opisu uzyskanych wyników; stosować procedury analizy i eksploracji danych; wykorzystywać procedury eksploracji danych do praktycznych zagadnień pozyskiwania wiedzy z danych;

*kompetencje społeczne:* (jest gotów do) samodzielnego poszerzania umiejętności tworzenia prezentacji otrzymywanych wyników; rozumienia istoty i ograniczenia danych liczbowych wykorzystywanych do badań; brania odpowiedzialności za przedstawioną interpretację wyników.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG6, SP\_P7S\_WG7, SP\_P7S\_WG8, SP\_P7S\_WK2, SP\_P7S\_UW1, SP\_P7S\_UW4, SP\_P7S\_UW5, SP\_P7S\_UK1, SP\_P7S\_UK2, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 3

### **4. Power BI**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest przygotowanie do sprawnego korzystania z Power BI – uczestnik będzie umiał przygotować dane, dokonywać na nich obliczeń, a także publikować gotowe raporty i dashboardy.

Treści merytoryczne:

- Importowanie danych
- Praca z modelem danych
- Wizualizacja danych
- Power BI Service.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) rozmaite techniki prezentacji danych; zalety i ograniczenia poznanych technik prezentacji; gotowe biblioteki do analizy danych;

*umiejętności:* (potrafi) dobrać odpowiednią technikę prezentacji do otrzymanych wyników; przygotować zestawienie danych w postaci graficznej, a następnie dokonać opisu uzyskanych wyników; stosować procedury analizy danych;

*kompetencje społeczne:* (jest gotów do) samodzielnego poszerzania umiejętności tworzenia prezentacji otrzymywanych wyników; rozumienia istoty i ograniczenia danych liczbowych wykorzystywanych do badań; brania odpowiedzialności za przedstawioną interpretację wyników.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG6, SP\_P7S\_WG7, SP\_P7S\_WG8, SP\_P7S\_WK2, SP\_P7S\_UW1, SP\_P7S\_UW4, SP\_P7S\_UW5, SP\_P7S\_UK1, SP\_P7S\_UK2, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 3

## **5. Bazy i źródła danych**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest przedstawienie podstaw baz danych i języków zapytań, architektury systemów baz danych oraz metod projektowania baz danych.

Treści merytoryczne:

- relacyjne bazy danych - język SQL,
- nierelacyjne bazy danych – MongoDB,
- zewnętrzne źródła danych - rest API,
- integracja Python z bazami danych.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) zasady projektowania baz danych; podstawowe własności języka zapytań SQL; zasady komunikacji języków programowania z serwerami baz danych;

*umiejętności:* (potrafi) zaprojektować i zaimplementować relacyjną bazę danych; budować oraz modyfikować konstrukcję zapytań do baz danych; zaimportować dane zewnętrzne do bazy;

*kompetencje społeczne:* (jest gotów do) uznania ograniczenia własnej wiedzy dotyczącej baz danych i rozumie potrzebę dalszego kształcenia; samodzielnego poszerzania umiejętności tworzenia i modyfikacji baz danych.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG3, SP\_P7S\_WG7, SP\_P7S\_WK2, SP\_P7S\_UW4, SP\_P7S\_UW6, SP\_P7S\_UW7, SP\_P7S\_UK1, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR2, SP\_P7S\_KR3, SP\_P7S\_KK1, SP\_P7S\_KO1

*Liczba ECTS:* 3

## **6. Statystyka i algebra w praktyce**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest przekazanie słuchaczom praktycznej wiedzy z zakresu statystyki i algebry z wykorzystaniem oprogramowania służącego do statystycznego opracowania danych.

Treści merytoryczne:

- odchylenie,
- korelacja,
- regresja liniowa,
- prawdopodobieństwo - rozkłady próbkowanie,
- testowanie hipotez, test  $\chi^2$ , ANOVA,
- układy równań - macierze i wektory.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe definicje i twierdzenia matematyki z zakresu algebry i kombinatoryki; pojęcia z zakresu wnioskowania statystycznego;

*umiejętności:* (potrafi) projektować i przeprowadzać badanie statystyczne zgodnie ze standardami wnioskowania statystycznego; zaimplementować język Python do wykonania obliczeń matematycznych; przetwarzać dane statystyczne dotyczące zjawisk masowych za pomocą wybranego oprogramowania;

*kompetencje społeczne:* (jest gotów do) uznania ograniczenia własnej wiedzy dotyczącej matematyki wyższej i rozumie potrzebę dalszego kształcenia; rozumienia istoty i ograniczenia danych liczbowych wykorzystywanych w badaniach statystycznych.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG1, SP\_P7S\_WG2, SP\_P7S\_WG5, SP\_P7S\_WG7, SP\_P7S\_UW1, SP\_P7S\_UW2, SP\_P7S\_UW3, SP\_P7S\_UW4, SP\_P7S\_UK1, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 1

## **7. Zaawansowany język Python**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest pogłębienie dotychczas zdobytej wiedzy na temat programowania w języku Python.

*Treści merytoryczne:*

- Virtualenv,
- programowanie obiektowe w Python,
- wyrażenia lambda,
- obsługa wyjątków,
- zaawansowane struktury danych,
- wyrażenia regularne,
- dekoratory,
- programowanie współbieżne i wielowątkowe.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) zaawansowane mechanizmy w Pythonie; strukturę języka;

*umiejętności:* (potrafi) używać zaawansowanych konstrukcji składniowych języka Python; modyfikować istniejące duże programy w Pythonie;

*kompetencje społeczne:* (jest gotów do) rozumienia potrzeby tworzenia czytelnych i wydajnych programów; samodzielnego wyszukiwania informacji w literaturze.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG7, SP\_P7S\_UW4, SP\_P7S\_UK1, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 3

## **8. Wprowadzenie do języka R**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest zapoznanie słuchacza z podstawami języka R i możliwościami jego implementacji.

Treści merytoryczne:

- instalacja środowiska,
- podstawowe elementy konstrukcyjne języka,
- podstawowe funkcje,
- tworzenie funkcji,
- instrukcje sterujące,
- import/eksport danych,
- projekt.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe typy i struktury danych języka R, symbole specjalne, podstawowe funkcje i instrukcje sterujące; zasady tworzenia własnych funkcji w języku R;

*umiejętności:* (potrafi) napisać i uruchomić program w języku R; korzystać z wybranych pakietów języka R; zaprezentować wyniki wykorzystując poznane oprogramowanie;

*kompetencje społeczne:* (jest gotów do) rozumienia, że biegłe posługiwanie się programem R wymaga ciągłego poznawania tego pakietu i doskonalenia warsztatu wyników; uznania faktu, że program R wraz z pakietami dodatkowymi jest nieustannie rozwijany i oferuje z czasem nowe możliwości.

*Symbolne efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG7, SP\_P7S\_UW4, SP\_P7S\_UK1, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 3

## **9. Uczenie maszynowe**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest wprowadzenie wybranych algorytmów stosowanych w robotyce mobilnej oraz IoE (Internecie Wszechrzeczy), w tym modelowanie mapy, lokalizacja na mapie, sterowanie serwo mechanizmami, śledzenie obiektów, planowanie ruchu, wygładzanie ruchu.

Treści merytoryczne:

- modelowanie procesów decyzyjnych, m.in. techniki:
  - metody regresji,
  - techniki klasyfikacji k-NN, naiwny klasyfikator Bayesa, SVM, klasyfikatory regułowe, sieci neuronowe (Deep Learning), drzewa decyzyjne - algorytm C4.5. Zastosowanie algorytmów rojowych i genetycznych. Ocena skuteczności modeli decyzyjnych w tym metoda walidacji krzyżowej standardowej oraz Monte Carlo, Bagging i Leave One Out.
  - techniki Ensemble (Random Forests, Bagging oraz Boosting)

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe pojęcia z dziedziny uczenia maszynowego; podstawowe metody klasyfikacji;

*umiejętności:* (potrafi) zastosować podejście uczenia maszynowego lub sztucznej sieci neuronowej do praktycznego problemu; konstruować model klasyfikujący dla zadanej bazy wiedzy; ocenić skuteczność budowanego modelu; wyprowadzać wnioski na podstawie eksperymentów; przygotować harmonogram dobierania metod w zależności od ich skuteczności; weryfikować postawione tezy badawcze i demonstrować rozwiązania;

*kompetencje społeczne:* (jest gotów do) doceniania znaczenia metod uczenia maszynowego we współczesnych metodach analizowania baz wiedzy; samodzielnego wyszukiwania informacji w literaturze.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG4, SP\_P7S\_WG5, SP\_P7S\_WG7, SP\_P7S\_UW4, SP\_P7S\_UW8,  
SP\_P7S\_UK1, SP\_P7S\_UK3, SP\_P7S\_UU1, SP\_P7S\_KR3, SP\_P7S\_KK1

*Liczba ECTS:* 4

## **10. Analiza dużych zbiorów danych I**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest zapoznanie słuchacza z nowymi technologiami informatycznymi służącymi przetwarzaniu dużych zbiorów danych oraz przedstawienie słuchaczom wybranych metod i algorytmów wydobywania wiedzy z dużych zbiorów danych.

*Treści merytoryczne:* 3 zagadnienia do wyboru:

- apache Hadoop,
- hadoop Map Reduce,
- apache Hive,
- apache Pig,
- apache Mahout.

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe metody, algorytmy i narzędzia pozyskiwania i integracji danych; możliwości efektywnego składowania dużych zbiorów danych;

*umiejętności:* (potrafi) korzystać z wybranych narzędzi przetwarzania dużych zbiorów danych w celu pozyskania z nich informacji i wiedzy; projektować i konstruować informatyczne środowiska gromadzenia dużych zbiorów danych; dokonywać analizy i prezentacji zgromadzonych danych i pozyskanych informacji na potrzeby praktyki w różnych dziedzinach;

*kompetencje społeczne:* (jest gotów do) rozumienia, że biegłe posługiwanie się umiejętnością analizy dużych zbiorów danych wymaga ciągłego doksztalcania się; brania odpowiedzialności za przedstawioną interpretację wyników.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG3, SP\_P7S\_WG6, SP\_P7S\_WG7, SP\_P7S\_WG8, SP\_P7S\_WK1,  
SP\_P7S\_WK2, SP\_P7S\_WK3,

SP\_P7S\_UW1, SP\_P7S\_UW4, SP\_P7S\_UW5, SP\_P7S\_UK1, SP\_P7S\_UK2, SP\_P7S\_UK3,  
SP\_P7S\_UO1, SP\_P7S\_UO2, SP\_P7S\_UU1, SP\_P7S\_UU2, SP\_P7S\_UU3, SP\_P7S\_KR1,  
SP\_P7S\_KR2, SP\_P7S\_KR3, SP\_P7S\_KK1, SP\_P7S\_KK2, SP\_P7S\_KO1, SP\_P7S\_KO2,  
SP\_P7S\_KO3

*Liczba ECTS:* 3

## **11. Analiza dużych zbiorów danych II**

*Cel kształcenia i treści merytoryczne:* celem zajęć jest zapoznanie słuchacza z nowymi technologiami informatycznymi służącymi przetwarzaniu dużych zbiorów danych oraz przedstawienie słuchaczom wybranych metod i algorytmów wydobywania wiedzy z dużych zbiorów danych.

*Treści merytoryczne:*

- wprowadzenie do Apache Spark,
- spark SQL,
- spark streaming,
- spark Mlib,

*Efekty uczenia się:*

*wiedza:* (zna i rozumie) podstawowe metody, algorytmy i narzędzia pozyskiwania i integracji danych; możliwości efektywnego składowania dużych zbiorów danych;

*umiejętności:* (potrafi) korzystać z wybranych narzędzi przetwarzania dużych zbiorów danych w celu pozyskania z nich informacji i wiedzy; projektować i konstruować informatyczne środowiska gromadzenia dużych zbiorów danych; dokonywać analizy i prezentacji zgromadzonych danych i pozyskanych informacji na potrzeby praktyki w różnych dziedzinach;

*kompetencje społeczne:* (jest gotów do) rozumienia, że biegłe posługiwanie się umiejętnością analizy dużych zbiorów danych wymaga ciągłego dokształcania się; brania odpowiedzialności za przedstawioną interpretację wyników.

*Symbole efektów uczenia się dla studiów podyplomowych:*

SP\_P7S\_WG2, SP\_P7S\_WG3, SP\_P7S\_WG6, SP\_P7S\_WG7, SP\_P7S\_WG8, SP\_P7S\_WK1, SP\_P7S\_WK2,

SP\_P7S\_UW1, SP\_P7S\_UW4, SP\_P7S\_UW5, SP\_P7S\_UK1, SP\_P7S\_UK2, SP\_P7S\_UK3, SP\_P7S\_UO1, SP\_P7S\_UO2, SP\_P7S\_UU1, SP\_P7S\_KR1, SP\_P7S\_KR2, SP\_P7S\_KR3, SP\_P7S\_KK1, SP\_P7S\_KK2, SP\_P7S\_KO1, SP\_P7S\_KO3

*Liczba ECTS:* 3

*Liczba ECTS:* 3

PLAN STUDIÓW PODYPLOMOWYCH

Nazwa studiów podyplomowych: „Data science w praktyce”

Wymiar kształcenia (sem.): dwa semestry

Liczba punktów ECTS konieczna do uzyskania kwalifikacji podyplomowych: 30

Lp.	Nazwa przedmiotu	Rodzaj i wymiar zajęć dydaktycznych				Forma zaliczenia przedmiotu/sposób weryfikacji efektów uczenia się	Punkty ECTS
		Wykłady (godz.)	Ćwiczenia (godz.)	Zajęcia teoretyczne (godz.)	Zajęcia praktyczne (godz.)		
Semestr I							
1	Wprowadzenie do języka Python	-	20	-	20	Zaliczenie	3
2	Warsztat badacza danych	-	15	-	15	Zaliczenie	1
3	Wizualizacja i eksploracja danych	-	20	-	20	Zaliczenie na ocenę	3
4	Power BI	-	15	-	15	Zaliczenie	3
5	Bazy i źródła danych	-	20	-	20	Zaliczenie na ocenę	3
6	Statystyka i algebra w praktyce	-	15	-	15	Zaliczenie	1
7	Zaawansowany język Python	-	20	-	20	Zaliczenie na ocenę	3
Semestr II							
8	Wprowadzenie do języka R	-	20	-	20	Zaliczenie	3
9	Uczenie maszynowe	-	30	-	30	Zaliczenie na ocenę	4
10	Analiza dużych zbiorów danych I	-	20	-	20	Zaliczenie	3
11	Analiza dużych zbiorów danych II	-	20	-	20	Zaliczenie na ocenę	3
Łączna liczba godzin		-	215	-	215	Łączna liczba punktów ECTS	30
		<b>215</b>		<b>215</b>			

Okres zaliczeniowy na studiach podyplomowych: 1 rok”