

TREŚCI KSZTAŁCENIA

Nazwa studiów podyplomowych: *analiza i inżynieria danych – data science*

Wymiar kształcenia (sem.): 2 semestry

CHARAKTERYSTYKA TREŚCI KSZTAŁCENIA

1. Wprowadzenie do data science

Cel kształcenia: zapoznanie z podstawami data science, nabycie umiejętności instalacji środowiska wybranego środowiska analitycznego, prezentacja najważniejszych pakietów i narzędzi.

Treści merytoryczne: Instalacja i konfiguracja środowiska pracy. Podstawowe elementy języka programowania: organizacja kodu, podstawowe typy danych, instrukcje warunkowe, pętle. Organizacja kodu: funkcje, moduły, pakiety oraz dokumentacja kodu. Obsługa narzędzia Jupyter Notebook. Podstawowe wykorzystanie pakietów pandas, matplotlib oraz seaborn w środowisku Jupyter Notebook. Język znaczników Markdown. Wykorzystanie systemu kontroli wersji Git.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): podstawowe funkcje wybranego środowiska programistycznego; metody pracy na danych; zastosowania i funkcjonalność wybranego oprogramowania; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): zainstalować i przygotować środowisko do pracy; napisać prosty program z wykorzystaniem języka Python; dobrać konstrukcje i struktury danych języka Python do realizacji zadanych operacji; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): ciągłego doksztalcania się; klarownego omówienia zastosowanych rozwiązań i technologii.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG9, SP_P6S_UW4, SP_P6S_UW9, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 4.

2. Wizualizacja danych i techniki data mining

Cel kształcenia: zapoznanie z metodami przeprowadzania eksploracji danych za pomocą poznanych metod data mining oraz zaznajomienie z metodami graficznej prezentacji danych.

Treści merytoryczne: Podstawowe metody statystyczne. Wczytywanie danych z różnych źródeł. Wybrane techniki data mining. Analiza sygnałów i szeregów czasowych. Podstawowe metody regresji liniowej i nieliniowej oraz prognozowania szeregów czasowych. Przetwarzanie danych tekstowych: normalizacja i wektoryzacja. Przetwarzanie zbiorów - zmiany formatu, brakujące wartości, przekształcanie, itp. Eksploracja danych - filtrowanie, sortowanie, agregacja (biblioteki numpy, pandas). Wizualizacja danych - przegląd najpopularniejszych bibliotek (matplotlib, seaborn, plotly, bokeh, altair).

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): sposób wykorzystania środowiska programistycznego języka Python wraz z wybranymi bibliotekami w przygotowaniu, obróbce i przeprowadzeniu analiz danych; możliwości aplikacyjne przedstawionych metod analitycznych i wizualizacji danych; rozmaite techniki prezentacji danych; specyfikę poszczególnych metod i modeli data mining; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): pozyskiwać i przetwarzać dane za pomocą wybranego środowiska programistycznego; wczytać dane do programu, określić jakość danych, dokonać podstawowych manipulacji na danych; przygotować dane do wybranej metody data mining; konstruować model klasyfikujący dla zadanej bazy wiedzy; stosować wybrane metody przeprowadzania eksploracji danych; przygotować zestawienie danych

w postaci tabelarycznej i graficznej, a następnie dokonać opisu uzyskanych wyników; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): uznania ograniczenia własnej wiedzy dotyczącej analizy danych i rozumie potrzebę dalszego kształcenia; brania odpowiedzialności za przedstawioną interpretację wyników; komunikowania ważnych wyników i osiągnięć społeczeństwu; przestrzegania przepisów dotyczących ochrony powierzonych do analizy danych.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG3, SP_P6S_WG5, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG8 SP_P6S_WG9, SP_P6S_UW4, SP_P6S_UW5, SP_P6S_UW9, SP_P6S_UK1, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 3.

3. Systemy baz danych

Cel kształcenia: zapoznanie z technologią systemów baz danych. Zapoznanie z architekturą systemów baz danych oraz metodami projektowania baz danych.

Treści merytoryczne: Relacyjne bazy danych - język SQL. Nierelacyjne bazy danych – Cassandra. Integracja Python z bazami danych. Programowanie baz danych PL/SQL.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): zasady projektowania baz danych; podstawowe własności języka zapytań SQL; zasady komunikacji języków programowania z serwerami baz danych; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): projektować i implementować relacyjną bazę danych; budować oraz modyfikować konstrukcję zapytań do baz danych; importować dane zewnętrzne do bazy; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

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

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG3, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG9, SP_P6S_WK2, SP_P6S_UW4, SP_P6S_UW6, SP_P6S_UW7, SP_P6S_UW9, SP_P6S_UK1, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 4.

4. Podstawy pakietu R

Cel kształcenia: zapoznanie z możliwościami wykorzystania pakietu R. Zapoznanie z metodami wizualizacji danych z wykorzystaniem języka R.

Treści merytoryczne: Instalacja środowiska. Podstawowe elementy konstrukcyjne języka. Podstawowe funkcje i ich tworzenie. Instrukcje sterujące. Import/eksport danych. Wizualizacja danych.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): podstawowe typy i struktury danych języka R, symbole specjalne, podstawowe funkcje i instrukcje sterujące; zasady tworzenia własnych funkcji w języku R; budowę i funkcjonalność pakietu do zdalnego nauczania.

Umiejętności (słuchacz potrafi): pisać i uruchamiać program w języku R; korzystać z wybranych pakietów języka R; prezentować wyniki wykorzystując poznane oprogramowanie; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): ciągłego poznawania pakietu R i doskonalenia warsztatu programistycznego; uznania faktu, że pakiet R wraz z pakietami dodatkowymi jest nieustannie rozwijany i oferuje z czasem nowe możliwości.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG9, SP_P6S_UW4, SP_P6S_UW9, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 3.

5. Power BI

Cel kształcenia i treści merytoryczne: zapoznanie z procesem przetwarzania i wizualizacji danych przy zastosowaniu Microsoft Power BI.

Treści merytoryczne: Importowanie i transformacja danych. Praca z modelem danych. Wizualizacja danych. Raportowanie.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): rozmaite techniki przetwarzania i wizualizacji danych; zalety i ograniczenia poznanych technik prezentacji; budowę i funkcjonalność pakietu wykorzystywanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): dobrać odpowiednią technikę prezentacji do otrzymanych wyników; przygotować zestawienie danych w postaci graficznej, a następnie dokonać opisu uzyskanych wyników; stosować procedury analizy danych; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz 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_P6S_WG2, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG8, SP_P6S_WG9, SP_P6S_UW4, SP_P6S_UW5, SP_P6S_UW9, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 2.

6. Zaawansowane programowanie w języku Python

Cel kształcenia: zapoznanie z zaawansowanymi technikami programowania w języku Python.

Treści merytoryczne: Programowanie zorientowane obiektowo. Moduły i pakiety. Obsługa plików. Dekoratory. Wyrażenia lambda. Usuwanie błędów, testowanie. Wyrażenia regularne.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): zaawansowane mechanizmy w języku Python; strukturę języka programowania; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): używać zaawansowanych konstrukcji składniowych języka Python; modyfikować istniejące duże programy w Pythonie; samodzielnie rozwiązywać problemy na każdym etapie przygotowania i realizacji programów i projektów w języku Python; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): tworzenia czytelnych i wydajnych programów; samodzielnego wyszukiwania informacji w literaturze; formułowania pytań, służących pogłębieniu własnej wiedzy dotyczącej wybranego języka programowania.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG9, SP_P6S_UW4, SP_P6S_UW9, SP_P6S_UK1, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 4.

7. Metody i techniki sztucznej inteligencji

Cel kształcenia: zapoznanie z metodami, modelami i narzędziami AI oraz zaprezentowanie zastosowań praktycznych wybranych algorytmów uczenia maszynowego.

Treści merytoryczne: Uczenie nadzorowane i regresja liniowa. Statystyki Bayesowskie. Drzewa decyzyjne. Uczenie nienadzorowane. Sieci neuronowe. Modele generatywne i autokodery. Algorytmy i metody uczenia modeli głębokich. Wybrane problemy klasyfikacji, detekcji, regresji.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): podstawowe pojęcia z zakresu uczenia maszynowego; podstawowe metody klasyfikacji; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): stosować 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; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): uznania ograniczenia własnej wiedzy, odczuwając potrzebę poszerzania swojej wiedzy i praktycznych umiejętności w zakresie metod uczenia maszynowego.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG1, SP_P6S_WG2, SP_P6S_WG4, SP_P6S_WG5, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG9, SP_P6S_UW1, SP_P6S_UW2, P_P6S_UW3, SP_P6S_UW4, SP_P6S_UW5, SP_P6S_UW8, SP_P6S_UU1, SP_P6S_KR1, SP_P6S_KR3, SP_P6S_KK1.

Liczba ECTS: 4.

8. Analiza danych w praktyce

Cel kształcenia: nabycie umiejętności praktycznego rozwiązywania wybranych problemów oraz raportowania otrzymanych wyników analiz.

Treści merytoryczne: Realizacja zadań praktycznych. Doskonalenie techniki prezentacji przeprowadzonej analizy.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): metody pozyskiwania, porządkowania i przechowywania zbiorów danych; podstawowe metody i narzędzia stosowane do transformacji danych; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): budować modele analizy danych w oparciu o różnorodne narzędzia; dobrać metodę analizy danych do wybranego problemu; dokonać analizy i prezentacji zgromadzonych danych i pozyskanych informacji; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): uznania ograniczenia własnej wiedzy dotyczącej analizy danych i rozumie potrzebę dalszego kształcenia; wzięcia odpowiedzialności za przedstawioną interpretację wyników, posiadając przy tym świadomość wpływu stosowanych metod na ich precyzję; komunikacji ważnych wyników i osiągnięć społeczeństwu; przestrzegania przepisów dotyczących ochrony powierzonych do analizy danych.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG4, SP_P6S_WG5, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG8, SP_P6S_WG9, SP_P6S_WK1, SP_P6S_WK2, SP_P6S_WK3, SP_P6S_UW1, SP_P6S_UW3, SP_P6S_UW4, SP_P6S_UW5, SP_P6S_UW9, SP_P6S_UK1, SP_P6S_UK2, SP_P6S_UK3, SP_P6S_UO1, SP_P6S_UO2, SP_P6S_UU1, SP_P6S_UU2, SP_P6S_KR1, SP_P6S_KR2, SP_P6S_KR3, SP_P6S_KK1, SP_P6S_KK2, SP_P6S_KO1, SP_P6S_KO2, SP_P6S_KO3.

Liczba ECTS: 2.

9. Inżynieria big data

Cel kształcenia: zapoznanie z technologiami przechowywania i przetwarzania dużych zbiorów danych. Prezentacja dostępnych narzędzi wykorzystywanych w inżynierii danych.

Treści merytoryczne: Wprowadzenie do inżynierii danych. Platforma Apache Hadoop. Podstawy Apache Spark. Batchowe i strumieniowe przetwarzanie danych. Budowa przepływów danych z użyciem Apache Airflow. Przetwarzanie danych w chmurze obliczeniowej. Infrastruktura big data.

Efekty uczenia się:

Wiedza (słuchacz zna i rozumie): metody pozyskiwania, porządkowania i przechowywania dużych zbiorów danych; podstawowe metody i narzędzia stosowane do transformacji dużych zbiorów danych; budowę i funkcjonalność pakietu używanego do zdalnego nauczania.

Umiejętności (słuchacz potrafi): budować modele analizy danych w oparciu o różnorodne narzędzia; dobrać metodę analizy dużych zbiorów danych do wybranego problemu; dokonać analizy i prezentacji zgromadzonych

danych i pozyskanych informacji; korzystać z nowoczesnych rozwiązań technologicznych dostarczonych przez wybraną platformę na potrzeby zdalnego uczenia się.

Kompetencje społeczne (słuchacz jest gotów do): uznania ograniczenia własnej wiedzy dotyczącej analizy danych i rozumie potrzebę dalszego kształcenia; wzięcia odpowiedzialności za przedstawioną interpretację wyników, posiadając przy tym świadomość wpływu stosowanych metod na ich precyzję; komunikacji ważnych wyników i osiągnięć społeczeństwu; przestrzegania przepisów dotyczących ochrony powierzonych do analizy danych.

Symbole efektów uczenia się dla studiów podyplomowych: SP_P6S_WG2, SP_P6S_WG3, SP_P6S_WG4, SP_P6S_WG6, SP_P6S_WG7, SP_P6S_WG8, SP_P6S_WG9, SP_P6S_WK1, SP_P6S_WK2, SP_P6S_WK3, SP_P6S_UW1, SP_P6S_UW4, SP_P6S_UW5, SP_P6S_UW6, SP_P6S_UW7, SP_P6S_UW9, SP_P6S_UK1, SP_P6S_UK2, SP_P6S_UK3, SP_P6S_UO1, SP_P6S_UO2, SP_P6S_UU1, SP_P6S_UU2, SP_P6S_KR1, SP_P6S_KR2, SP_P6S_KR3, SP_P6S_KK1, SP_P6S_KK2, SP_P6S_KO1, SP_P6S_KO2, SP_P6S_KO3.

Liczba ECTS: 4.

PLAN STUDIÓW PODYPLOMOWYCH

Nazwa studiów podyplomowych: *analiza i inżynieria danych – data science*

Wymiar kształcenia (sem.): dwa semestry

Liczba punktów ECTS konieczna do uzyskania kwalifikacji podyplomowych: 30

Lp.	Nazwa przedmiotu/zajęć	Forma i wymiar zajęć dydaktycznych			Forma zaliczenia przedmiotu/sposób weryfikacji efektów uczenia się	Punkty ECTS
		Rodzaj zajęć	Zajęcia teoretyczne (godz.)	Zajęcia praktyczne (godz.)		
Semestr I						
1	Wprowadzenie do data science	ćwiczenia	-	30	zal./ aktywność na zajęciach, mini-projekt	4
2	Wizualizacja danych i techniki data mining	ćwiczenia	-	20	zal. oc./ aktywność na zajęciach, mini-projekt	3
3	Systemy baz danych	ćwiczenia	-	30	zal. oc. / aktywność na zajęciach, mini-projekt	4
4	Podstawy pakietu R	ćwiczenia	-	20	zal./ aktywność na zajęciach, mini-projekt	3
5	Power BI	ćwiczenia	-	15	zal./ aktywność na zajęciach, mini-projekt	2
Semestr II						
6	Zaawansowane programowanie w języku Python	ćwiczenia	-	25	zal. oc./ aktywność na zajęciach, mini-projekt	4
7	Metody i techniki sztucznej inteligencji	ćwiczenia	-	30	zal. oc./ aktywność na zajęciach, mini-projekt	4
8	Analiza danych w praktyce	ćwiczenia	-	15	zal. oc./ aktywność na zajęciach, mini-projekt	2
9	Inżynieria big data	ćwiczenia	-	30	zal. oc./ aktywność na zajęciach, mini-projekt	4
Łączna liczba godzin:		x	-	215	Łączna liczba punktów ECTS:	30
				215		

Okres zaliczeniowy na studiach podyplomowych: 1 rok.

Objaśnienia:

¹ Wykłady/ćwiczenia.

² Symbole formy zaliczenia: zal. – zaliczenie bez oceny, zal. oc. – zaliczenie na ocenę, egz. – egzamin.