



Uniwersytet
Warmińsko-
Mazurski w
Olsztynie



Studia podyplomowe: Zaawansowane metody analizy danych i Data Mining w biznesie

Numer usługi 2024/05/09/12141/2145292

📍 zdalna w czasie rzeczywistym

📚 Studia podyplomowe

🕒 210 h

📅 26.10.2024 do 08.06.2025

6 900,00 PLN brutto

6 900,00 PLN netto

32,86 PLN brutto/h

32,86 PLN netto/h

Informacje podstawowe

Kategoria	Informatyka i telekomunikacja / Programowanie
Sposób dofinansowania	wsparcie dla osób indywidualnych wsparcie dla pracodawców i ich pracowników
Grupa docelowa usługi	Studia skierowane są do przedsiębiorców i pracowników, w szczególności: banków, instytucji ubezpieczeniowych, przedsiębiorstw handlowych, produkcyjnych, ośrodków przetwarzania informacji oraz ośrodków badania opinii społecznej, firm prowadzących badania kliniczne, instytucji administracji państwowej samorządowej.
Minimalna liczba uczestników	10
Maksymalna liczba uczestników	15
Data zakończenia rekrutacji	25-10-2024
Forma prowadzenia usługi	zdalna w czasie rzeczywistym
Liczba godzin usługi	210
Podstawa uzyskania wpisu do BUR	art. 163 ust. 1 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (t.j. Dz. U. z 2023 r. poz. 742, z późn. zm.)
Zakres uprawnień	Studia podyplomowe

Cel

Cel edukacyjny

Celem studiów jest zapoznanie słuchaczy z najbardziej popularnymi technikami, algorytmami, narzędziami i oprogramowaniem stosowanym w Data Science oraz w Big Data.

Usługa Zaawansowane metody analizy danych i Data Mining w biznesie przygotowuje do samodzielnego wykonywania przetwarzania, analizy i wizualizacji danych, tworzenia modeli predykcyjnych oraz analizy dużych zbiorów danych.

Efekty uczenia się oraz kryteria weryfikacji ich osiągnięcia i Metody walidacji

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
Programuje w wybranym języku programowania	Charakteryzuje metody i techniki programowania; Stosuje techniki obliczeniowe oraz techniki programowania, wspomagające pracę analityka; Dokonuje ilustracji obliczeń symbolicznych za pomocą pakietów oprogramowania; Projektuje i uzasadnia poprawność działania programu z uwzględnieniem złożoności algorytmów i zapisuje go w języku wysokiego poziomu; Rozumie ograniczenia własnej wiedzy i umiejętności, odczuwa potrzebę dalszego kształcenia, w tym zdobywania wiedzy pozadzielinowej	Obserwacja w warunkach symulowanych
		Obserwacja w warunkach rzeczywistych
Tworzy i zarządza systemami baz danych	Wymienia zasady dotyczące projektowania, tworzenia i zarządzania systemami baz danych; Posługuje się właściwie dobranymi środowiskami programistycznymi do projektowania, tworzenia, modyfikacji i zarządzania bazami danych; Analizuje złożoność struktur i baz danych, proponuje stosowne procedury, ocenia ich poprawność oraz implementuje je w wybranym języku programowania; Ocenia możliwości wykorzystania dotychczasowych osiągnięć technologii w swoim zawodzie	Obserwacja w warunkach rzeczywistych
		Obserwacja w warunkach symulowanych

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
<p>Stosuje metody i techniki sztucznej inteligencji; Przygotowuje i buduje model uczenia maszynowego</p>	<p>Wymienia zasady dotyczące reprezentowania wiedzy oraz mechanizmów klasyfikujących; Podaje przykłady ilustrujące konkretne pojęcia matematyczne; Stosuje techniki obliczeniowe oraz techniki programowania, wspomagające pracę analityka; Dokonuje ilustracji obliczeń symbolicznych za pomocą pakietów oprogramowania; podaje różne przykłady rozkładów prawdopodobieństwa dyskretnych i ciągłych i omawia wybrane eksperymenty losowe oraz modele matematyczne, w jakich te rozkłady występują; Posługuje się charakterystykami statystycznymi populacji i ich odpowiednikami próbkowymi; Projektuje i uzasadnia poprawność działania programu z uwzględnieniem złożoności algorytmów i zapisuje go w języku wysokiego poziomu; Implementuje poznane algorytmy w zakresie zagadnień związanych z wizualizacją komputerową;</p>	<p>Obserwacja w warunkach rzeczywistych</p> <p>Obserwacja w warunkach symulowanych</p>
<p>Stosuje wybrane modele matematyczne i techniki eksploracji do rozwiązywania wybranych zadań analizy danych; Dokonuje wizualizacji danych za pomocą wybranego oprogramowania</p>	<p>Definiuje pojęcia matematyczne z zakresu algebry i statystyki; Podaje przykłady ilustrujące konkretne pojęcia matematyczne; Definiuje pojęcia dotyczące wizualizacji danych na komputerze; Dobiera odpowiedni model statystyczny do analizy danych oraz implementuje go w praktyce przy pomocy oprogramowania; Podaje różne przykłady rozkładów prawdopodobieństwa dyskretnych i ciągłych i omawia wybrane eksperymenty losowe oraz modele matematyczne, w jakich te rozkłady występują; Posługuje się charakterystykami statystycznymi populacji i ich odpowiednikami próbkowymi; Implementuje poznane algorytmy w zakresie zagadnień związanych z wizualizacją komputerową; Rozumie ograniczenia własnej wiedzy i umiejętności, odczuwa potrzebę dalszego kształcenia, w tym zdobywania wiedzy pozadziędzinowej</p>	<p>Obserwacja w warunkach rzeczywistych</p> <p>Obserwacja w warunkach symulowanych</p>

Efekty uczenia się	Kryteria weryfikacji	Metoda walidacji
<p>Korzysta z wybranych technik i narzędzi przetwarzania dużych zbiorów danych w celu pozyskania z nich informacji i wiedzy;</p>	<p>Prezentuje techniki obliczeniowe oraz techniki programowania, wspomagające pracę analityka; Dobiera odpowiedni model statystyczny do analizy danych oraz implementuje go w praktyce przy pomocy oprogramowania; Pracuje z odbiorcami tworzonych rozwiązań informatycznych i analitycznych, aktywnie uczestniczy w dyskusji o potrzebach, możliwych rozwiązaniach i zasadach pozyskania, przetwarzania danych oraz ich wykorzystania;</p>	<p>Obserwacja w warunkach rzeczywistych</p>
	<p>Ocenia możliwości wykorzystania dotychczasowych osiągnięć technologii w swoim zawodzie; Zachowuje się w sposób profesjonalny, przestrzega zasad etyki zawodowej i poszanowania różnorodności poglądów; uznaje zawód informatyka oraz analityka danych jako rolę społeczną i uwzględnia problemy związane z poufnością danych; Komunikuje się ze specjalistami w swojej dziedzinie</p>	<p>Obserwacja w warunkach symulowanych</p>

Kwalifikacje

Kompetencje

Usługa prowadzi do nabycia kompetencji.

Warunki uznania kompetencji

Pytanie 1. Czy dokument potwierdzający uzyskanie kompetencji zawiera opis efektów uczenia się?

Do świadectwa ukończenia studiów podyplomowych absolwent otrzymuje jako załącznik zaświadczenie zawierające opis efektów uczenia się w zakresie wiedzy, umiejętności oraz kompetencji społecznych. Po ukończeniu studiów uzyskuje kwalifikacje cząstkowe na poziomie 7 PRK

Pytanie 2. Czy dokument potwierdza, że walidacja została przeprowadzona w oparciu o zdefiniowane w efektach uczenia się kryteria ich weryfikacji?

Do świadectwa ukończenia studiów podyplomowych absolwent otrzymuje jako załącznik zaświadczenie zawierające informację o przeprowadzeniu walidacji w oparciu o zdefiniowane w efektach uczenia się kryteria ich weryfikacji.

Pytanie 3. Czy dokument potwierdza zastosowanie rozwiązań zapewniających rozdzielenie procesów kształcenia i szkolenia od walidacji?

Do świadectwa ukończenia studiów podyplomowych absolwent otrzymuje jako załącznik zaświadczenie zawierające informację o rozdzieleniu procesu kształcenia od walidacji oraz nazwisko osoby przeprowadzającej walidację studiów.

Program

110 godz.dyd. w semestrze zimowym, 100 godz.dyd. w semestrze letnim* (godzina dydaktyczna = 45 minut).

Forma wszystkich zajęć ćwiczenia laboratoryjne

Zjazdy sobotnio-niedzielne online w czasie rzeczywistym na platformie Google Workspace.

Ramowy program usługi

Wprowadzenie do narzędzi analitycznych (35 ćw.)

Instalacja i konfiguracja środowiska pracy.

Podstawowe elementy języka Python: organizacja kodu, podstawowe typy danych, instrukcje warunkowe, pętle.

Organizacja kodu: funkcje, moduły, pakiety oraz dokumentacja kodu.

Wprowadzenie do narzędzia Jupyter Notebook.

Podstawowe wykorzystanie pakietów pandas, matplotlib oraz seaborn w środowisku Jupyter Notebook.

Język znaczników Markdown.

Wykorzystanie systemu kontroli wersji Git

Eksploracja i wizualizacja danych (20 ćw.)

Wczytywanie danych z różnych źródeł.

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).

Bazy danych (35 ćw.)

Relacyjne bazy danych - język SQL.

Nierelacyjne bazy danych – Cassandra.

Integracja Python z bazami danych.

Programowanie baz danych PL/SQL.

Elementy data mining (20 ćw.)

1. Importowanie danych
2. Praca z modelem danych
3. Wizualizacja danych
4. Power BI Service.

Bazy i źródła danych (20 ćw.)

Wprowadzenie do standardu CRISP-DM.

Podstawowe metody statystyczne:

- badanie rozkładu klas decyzyjnych,
- częstości wartości,
- miary rozproszenia, tendencji centralnej,
- obliczanie korelacji pomiędzy zmiennymi oraz wpływu atrybutów warunkowych na klasę decyzyjną (positive ratio).

Wybrane techniki pracy z danymi:

- standaryzacja,
- normalizacja,
- uzupełnianie uszkodzonych danych,
- konwersja wartości symbolicznych do numerycznych.

Analiza sygnałów i szeregów czasowych

Podstawowe metody regresji liniowej i nieliniowej oraz prognozowania szeregów czasowych.

Przetwarzanie danych tekstowych: normalizacja i wektoryzacja.

Zastosowanie języka Python do eksploracji, analizy i przetwarzania danych

Programowanie zaawansowane (20 ćw.)

Moduły i pakiety.

Obsługa plików.

Dekoratory.

Wyrażenia lambda.
Usuwanie błędów, testowanie.
Wyrażenia regularne.

Wprowadzenie do uczenia maszynowego (30 ćw.)

Uczenie nadzorowane i regresja liniowa.
Statystyki Bayesowskie.
Drzewa decyzyjne.
Uczenie nienadzorowane.
Sieci neuronowe.
Modele generatywne i autokodery

Uczenie maszynowe w praktyce (20 ćw.)

Algorytmy i metody uczenia modeli głębokich.
Wybrane problemy klasyfikacji, detekcji, regresji.

Wprowadzenie do Big Data (30 ćw.)

Wprowadzenie do Big Data.
Architektura i technologie Big Data.
Platforma Apache Hadoop.
Podstawy Apache Spark.
Batchowe i strumieniowe przetwarzanie danych.
Budowa przepływów danych z użyciem Apache Airflow.
Uczenie maszynowe w Big Data

Harmonogram

Liczba przedmiotów/zajęć: 0

Przedmiot / temat zajęć	Data realizacji zajęć	Godzina rozpoczęcia	Godzina zakończenia	Liczba godzin
Brak wyników.				

Cennik

Cennik

Rodzaj ceny	Cena
Koszt przypadający na 1 uczestnika brutto	6 900,00 PLN
Koszt przypadający na 1 uczestnika netto	6 900,00 PLN
Koszt osobogodziny brutto	32,86 PLN
Koszt osobogodziny netto	32,86 PLN

Prowadzący

Liczba prowadzących: 5



1 z 5

Krzysztof Ropiak

mgr matematyki, asystent na Wydziale Matematyki i Informatyki UWM w Olsztynie. Prowadzi zajęcia na I i II stopniu na kierunku informatyka od 2010r.



2 z 5

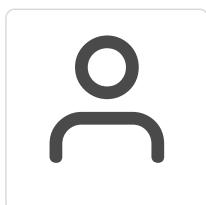
Tomasz Krzywicki

Obszar specjalizacji: mgr inż. informatyki

Doświadczenie zawodowe: asystent WMiI UWM w Olsztynie

Doświadczenie w świadczeniu tego typu usług: prowadzenie zajęć na studiach I i II stopnia na kierunku informatyka (od 2019r.), prowadzenie zajęć na studiach podyplomowych Data science w praktyce, Data science w Python, Zaawansowane technologie informatyczne

Wykształcenie: mgr inż. informatyki

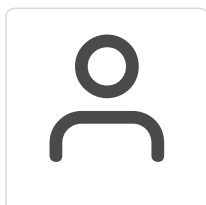


3 z 5

Paweł Drozda

Dr informatyki, adiunkt na Wydziale Matematyki i Informatyki UWM w Olsztynie. Prowadzi zajęcia na studiach I i II stopnia na kierunku informatyka (od 2005r.), prowadzi zajęcia na studiach podyplomowych Data science w praktyce.

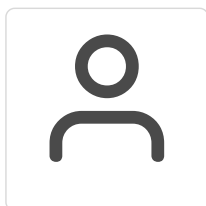
Prowadzenie szkoleń Statystyka, bazy danych, MS Office, analiza danych.



4 z 5

Michał Korpusik

Absolwent kierunków matematyka oraz informatyka. Programista, analityk z wieloletnim doświadczeniem w zakresie wizualizacji i eksploracji danych.



5 z 5

Paweł Procaj

Absolwent kierunku informatyka Politechniki Gdańskiej. Właściciel firmy informatycznej, analityk i programista z wieloletnim doświadczeniem.

Informacje dodatkowe

Informacje o materiałach dla uczestników usługi

materiały dydaktyczne udostępniane w postaci elektronicznej:

- skrypty z opisem teorii
- prezentacje w formie slajdów
- zadania z rozwiązaniami
- kody skryptów w językach programowania

Warunki uczestnictwa

Warunkiem ubiegania się o przyjęcia na studia podyplomowe jest posiadanie dyplomu ukończenia studiów co najmniej pierwszego stopnia dowolnego kierunku.

Informacje dodatkowe

110 godz.dyd. w semestrze zimowym, 100 godz.dyd. w semestrze letnim*

*godzina dydaktyczna = 45 minut

Kadra:

dr hab. Adam Lecko, prof. UWM - Kierownik studiów

dr Paweł Drozda (Katedra Metod Matematycznych Informatyki WMil UWM w Olsztynie)

mgr Krzysztof Ropiak (Katedra Metod Matematycznych Informatyki WMil UWM w Olsztynie)

mgr inż. Tomasz Krzywicki (Katedra Metod Matematycznych Informatyki WMil UWM w Olsztynie)

mgr inż. Michał Korpusik (IT Services Michał Korpusik)

mgr inż. Paweł Procaj (ProLogic Labs. IT Consulting Paweł Procaj)

Usługa jest rejestrowana na potrzeby usługodawcy i korzystającego z usługi jak również na potrzeby monitoringu, kontroli oraz w celu utrwalenia efektów kształcenia

Warunki techniczne

1. platforma / rodzaj komunikatora, za pośrednictwem której prowadzona będzie usługa

Przeglądarka internetowa Chrome lub Firefox w aktualnych wersjach dostępnych na stronach internetowych producenta / pakiet Google Workspace (Hangouts Meet)

1. minimalne wymagania sprzętowe, jakie musi spełniać komputer Uczestnika lub inne urządzenie do zdalnej komunikacji,

Parametry sprzętowe umożliwiające płynne działanie systemu operacyjnego minimalnie 4GB pamięci RAM, procesor intel core i3 lub odpowiednik. System operacyjny Windows 8.1 lub wyższy, MacOS i Linux w aktualny wersjach.

Komputer Uczestnika musi posiadać lub mieć podłączone sprawny mikrofon i kamerę. Uczestnik spotkania zobowiązany jest do wcześniejszej weryfikacji sprawności oraz konfiguracji mikrofonu i kamery.

c) minimalne wymagania dotyczące parametrów łącza sieciowego, jakim musi dysponować Uczestnik,

10 Mb/s – minimalna prędkość internetu

d) niezbędne oprogramowanie umożliwiające Uczestnikom dostęp do prezentowanych treści i materiałów,

System operacyjny Windows 7/8/10 w przypadku PC/laptopów, Android w wersji 5.0 w przypadku tabletów,

System operacyjny, przeglądarka internetowa, przeglądarka plików PDF oraz zgodnie z pkt 1 oraz łącze internetowe o parametrach opisanych w pkt 3

Notepad++

JAVA JDK

Pycharm

Powerbi desktop

Pyscripter/spider,

biblioteki do analizy danych

Git SCM

Mongo DB

VirtualBox

SQL Developer

Putty

XAMPP

Cassandra

Python (pakiety numpy, scipy, pandas, matplotlib, statsmodels, MySQLdb, psycopg2, os, xlswriter, cassandra)

Microsoft Office / Libre Office

e)okres ważności linku umożliwiającego uczestnictwo w spotkaniu on-line.

W okresie trwania szkolenia. Materiały związane z przygotowaniem środowiska do szkolenia będą dostępne bez ograniczeń czasowych.

Kontakt



Kinga Lecko

E-mail kinga@matman.uwm.edu.pl

Telefon (+48) 664 315 525