

Analiza danych w języku Python z elementami Big Data

STUDIA PODYPLOMOWE



Program zajęć

9

178

11

2

Liczba miesięcy nauki Liczba godzin zajęć Liczba zjazdów Liczba semestrów

Środowisko analityczne i podstawy pracy z Pythonem (20 godz.)

- Wprowadzenie do pracy analityka danych z wykorzystaniem języka Python
- Przygotowanie środowiska pracy
- Podstawowe konstrukcje języka Python
- Typy danych i podstawowe operacje
- Pisanie prostych programów i notatników
- Wykorzystanie podstawowych bibliotek do analizy danych

Pozyskiwanie danych w Pythonie (20 godz.)

- Pozyskiwanie danych do analizy z różnych źródeł
- Praca z plikami
- Pobieranie danych z zewnętrznych interfejsów API
- Podstawy pracy z relacyjnymi bazami danych
- Wczytywanie i przygotowanie danych do dalszego przetwarzania

Przetwarzanie i przygotowanie danych w Pythonie (20 godz.)

- Czyszczenie danych
- Obsługa braków i błędów
- Transformacja danych
- Łączenie i porządkowanie zbiorów danych
- Przygotowanie spójnych zbiorów danych do analizy

Eksploracyjna analiza danych w Pythonie (20 godz.)

- Podstawowe metody statystyki opisowej
- Rozkłady danych
- Zależności i korelacje między zmiennymi
- Eksploracja danych w Pythonie
- Interpretacja wyników analizy



- Formułowanie wniosków na podstawie danych

Wizualizacja danych w Pythonie (20 godz.)

- Narzędzia komunikacji wyników analizy
- Tworzenie czytelnych wykresów
- Dobór formy wizualizacji do rodzaju danych
- Prezentacja wyników analizy
- Przygotowanie wizualizacji wspierających interpretację danych

Modelowanie danych w Pythonie (20 godz.)

- Wprowadzenie do budowy modeli analitycznych w Pythonie
- Regresja
- Klasyfikacja
- Klasteryzacja
- Przygotowanie danych do modelowania
- Ocena jakości modeli
- Budowa prostych modeli analitycznych

Wprowadzenie do Big Data i hurtowni danych (20 godz.)

- Podstawowe pojęcia związane z Big Data
- Koncepcja hurtowni danych
- Modele OLTP i OLAP
- Proces ETL/ELT
- Miejsce rozwiązań analitycznych w architekturze danych
- Znaczenie nowoczesnych ekosystemów danych

Jakość i wspomaganie analizy danych narzędziami AI (20 godz.)

- Zapewnienie jakości i wiarygodności analiz danych
- Weryfikacja danych i wyników
- Ograniczanie błędów analitycznych
- Organizacja kodu analitycznego
- Wykorzystanie narzędzi AI do wspomagania pracy analityka



- Weryfikacja i ocena wyników generowanych przez AI

Konsultacje merytoryczne (16 godz.)

- Konsultacje do każdego modułu
- Sesje pytań i odpowiedzi
- Omówienie materiału po każdym module
- Wsparcie w realizacji zadań i przygotowaniu do zaliczenia

Egzamin (2 godz.)

- Egzamin (2 godz.)