

# Big Data. Inżynieria danych w AI

STUDIA PODYPŁOMOWE



## Program studiów

**9**

**194**

**10**

**2**

Liczba miesięcy nauki Liczba godzin zajęć Liczba zjazdów Liczba semestrów

### Przetwarzanie danych dla AI (8 godz.)

- Feature engineering (2 godz.)
- Data pipelines z użyciem Apache Airflow (4 godz.)
- Automatyzacja ETL pod modele ML (2 godz.)

### Modelowanie predykcyjne i ewaluacja modeli (20 godz.)

- Modelowanie z nadzorem (metody klasyfikacji, regresji), dobór parametrów oraz zmiennych w modelach uczelnia maszynowego (20 godz.)

### AI w Big data - zastosowania praktyczne (14 godz.)

- Wykorzystanie Spark MLlib (6 godz.)
- AI w strumieniowym przetwarzaniu danych (Kafka i ML) (8 godz.)

### Mlops i wdrażanie modeli AI (14 godz.)

- CI/CD dla modeli ML (3 godz.)
- Monitorowanie i retraining (3 godz.)
- Narzędzia: MLflow, Docker, Kubernetes (8 godz.)

### Systemy baz danych. Podstawy SQL-a (24 godz.)

- Diagramy ERD, Normalizacja, MS SQL, DDL, SQL DML (24 godz.)

### Programowanie w języku Python (24 godz.)

- Składnia, tablice, funkcje, Pandas, statystyka w Pythonie (24 godz.)

### Programowanie obiektowe w języku Python (12 godz.)

- Atrybuty, klasy, konstruktor, metody, dziedziczenie, „metody magiczne” (12 godz.)

### Nosql (Microsoft Azure) (20 godz.)

- Podstawowe koncepcje baz NoSQL - HBase, Cassandra, Impala, Neo4j (20 godz.)



## **Zaawansowane bazy danych i hurtowanie danych (24 godz.)**

- Zaawansowane aspekty języka SQL i TSQL; Koncepcje modelowania hurtowni danych (ROLAP, MOLAP, HOLAP); Technologie ETL/ELT; Elementy prezentacji danych np. Power BI (24 godz.)

## **Narzędzia Big Data (Microsoft Azure) (24 godz.)**

- Apache Hadoop & Apache Spark (24 godz.)

## **Projekt (8 godz.)**

- Seminarium projektowe (8 godz.)

## **Forma zaliczenia (2 godz.)**

- Test (1 godz.)
- Egzamin końcowy (1 godz.)