*Załącznik nr 1.5 do Zarządzenia Rektora UR nr 7/2023*

**SYLABUS**

**dotyczy cyklu kształcenia** 2024-2027

*(skrajne daty*)

**Rok akademicki 2026/2027**

1. Podstawowe informacje o przedmiocie

|  |  |
| --- | --- |
| Nazwa przedmiotu | Data mining |
| Kod przedmiotu\* | S1S[5]C\_01 |
| nazwa jednostki prowadzącej kierunek | Kolegium Nauk Społecznych |
| Nazwa jednostki realizującej przedmiot | Instytut Nauk Socjologicznych |
| Kierunek studiów | Socjologia |
| Poziom studiów | I stopnia |
| Profil | ogólnoakademicki |
| Forma studiów | stacjonarne |
| Rok i semestr/y studiów | Rok 3, semestr V |
| Rodzaj przedmiotu | specjalnościowy |
| Język wykładowy | jęz. polski |
| Koordynator | Katarzyna Garwol |
| Imię i nazwisko osoby prowadzącej / osób prowadzących | Katarzyna Garwol |

\* *-opcjonalni*e, *zgodnie z ustaleniami w Jednostce*

1.1.Formy zajęć dydaktycznych, wymiar godzin i punktów ECTS

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Semestr  (nr) | Wykł. | Ćw. | Konw. | Lab. | Sem. | ZP | Prakt. | Inne (zajęcia warsztatowe) | **Liczba pkt. ECTS** |
| V | - | - | 60 | - | - | - | - | - | 5 |

1.2. Sposób realizacji zajęć

X zajęcia w formie tradycyjnej

☐ zajęcia realizowane z wykorzystaniem metod i technik kształcenia na odległość

1.3 Forma zaliczenia przedmiotu (z toku) (egzamin, zaliczenie z oceną, zaliczenie bez oceny)

Zaliczenie z oceną

2.Wymagania wstępne

|  |
| --- |
| Znajomość podstaw statystyki. |

3. cele, efekty uczenia się, treści Programowe i stosowane metody Dydaktyczne

3.1 Cele przedmiotu

|  |  |
| --- | --- |
| C1 | Zapoznanie studentów socjologii z technikami data miningu oraz ich zastosowaniem w analizie danych społecznych. |
| C2 | Zapoznanie studentów z możliwościami jakie daje ekstrakcja użytecznej wiedzy z danych, która może być wykorzystana do podejmowana decyzji biznesowych, prognozowania trendów, odkrywania ukrytych wzorów itp. |
| C3 | Zapoznanie studentów z oprogramowaniem do analizy i eksploracji danych. |

**3.2 Efekty uczenia się dla przedmiotu**

|  |  |  |
| --- | --- | --- |
| EK (efekt uczenia się) | Treść efektu uczenia się zdefiniowanego dla przedmiotu | Odniesienie do efektów kierunkowych [[1]](#footnote-1) |
| EK\_01 | Student zna i rozumie w zaawansowanym stopniu metody i narzędzia oraz techniki pozyskiwania danych z wykorzystaniem narzędzi predefiniowanych do data mining, które pozwalają opisywać struktury i instytucje społeczne oraz procesy w nich i między nimi zachodzące. | KW\_07 |
| EK\_02 | Student potrafi prawidłowo interpretować zjawiska społeczne w zakresie socjologii oraz samodzielnie interpretować teorie socjologiczne, pozyskiwać dane do analizowania konkretnych procesów i zjawisk w zakresie socjologii, narzędziami do data mining. | KU\_01, KU\_02 |
| EK\_03 | Student potrafi analizować przyczyny i przebieg konkretnych zjawisk i procesów społecznych oraz prognozować te procesy i zjawiska z wykorzystaniem metod i narzędzi Data mining. | KU\_03, KU\_04 |
| EK\_04 | Student potrafi samodzielnie analizować zjawiska społeczne przy użyciu oprogramowania do data mining a także uzupełniać i doskonalić swoją wiedzę i umiejętności w tym zakresie. | KU\_08, KU\_14 |
| EK\_05 | Student jest gotów do uznania znaczenia wiedzy i krytycznej oceny posiadanych informacji oraz przedsiębiorczego myślenia i działania. | KK\_04, KK\_o5 |

**3.3 Treści programowe**

1. Problematyka ćwiczeń audytoryjnych, konwersatoryjnych, laboratoryjnych, zajęć praktycznych

|  |
| --- |
| Treści merytoryczne |
| Wprowadzenie do Data miningu |
| Podstawowe pojęcia związane z eksploracją danych |
| Źródła danych w socjologii |
| Przetwarzanie i wizualizacja danych |
| Przetwarzanie i czyszczenie danych: usuwanie brakujących wartości, normalizacja danych tekstowych, usuwanie duplikatów, przetwarzanie danych tekstowych i obrazowych w kontekście analizy społecznej, wizualizacja danych społecznych |
| Klasyfikacja danych: regresja logistyczna, drzewa decyzyjne, metoda k-najbliższych sąsiadów (k-NN), maszyny wektorów nośnych (SVM) |
| Grupowanie danych: k-means, analiza hierarchiczna skupień, reguły asocjacyjne, wykorzystywanie grupowania do identyfikacji segmentów społecznych i analizy związków między nimi. |
| Metody uczenia maszynowego: podstawy uczenia nadzorowanego i niedozorowanego, techniki regresji i klasyfikacji, sieci neuronowe (wprowadzenie i zastosowanie) |
| Text mining |
| Narzędzia do Data miningu: przegląd najpopularniejszych narzędzi i języków programowania do eksploracji danych (np. Python, pakiet R, Microsoft Power BI), praktyczne ćwiczenia z wykorzystaniem wybranego narzędzia. |
| Praktyczne zastosowania Data miningu: studia przypadków dla zastosowań Data miningu w różnych dziedzinach (np. biznes, marketing, medycyna, finanse), projekt badawczy (analiza danych z rzeczywistego zbioru danych i wyciąganie wniosków) |
| Etyka i bezpieczeństwo danych |

3.4 Metody dydaktyczne

Konwersatorium: wykład problemowy z prezentacją multimedialną, analiza wybranych studiów przypadku, ćwiczenia praktyczne z wykorzystaniem narzędzi do analizy danych, projekt badawczy.

4. METODY I KRYTERIA OCENY

4.1 Sposoby weryfikacji efektów uczenia się

|  |  |  |
| --- | --- | --- |
| Symbol efektu | Metody oceny efektów uczenia się  (np.: kolokwium, egzamin ustny, egzamin pisemny, projekt, sprawozdanie, obserwacja w trakcie zajęć) | Forma zajęć dydaktycznych  (w, ćw, …) |
| ek\_ 01 | Projekt badawczy, ćwiczenia praktyczne, obserwacja na zajęciach | Konw. |
| Ek\_ 02 | Projekt badawczy, ćwiczenia praktyczne, obserwacja na zajęciach | Konw. |
| EK\_03 | Projekt badawczy, ćwiczenia praktyczne, obserwacja na zajęciach | Konw. |
| EK\_04 | Projekt badawczy, ćwiczenia praktyczne, obserwacja na zajęciach | Konw. |
| EK\_05 | Projekt badawczy, ćwiczenia praktyczne, obserwacja na zajęciach | Konw. |

4.2 Warunki zaliczenia przedmiotu (kryteria oceniania)

|  |
| --- |
| Warunkiem zaliczenia przedmiotu jest: |

**5. CAŁKOWITY NAKŁAD PRACY STUDENTA POTRZEBNY DO OSIĄGNIĘCIA ZAŁOŻONYCH EFEKTÓW W GODZINACH ORAZ PUNKTACH ECTS**

|  |  |
| --- | --- |
| **Forma aktywności** | **Średnia liczba godzin na zrealizowanie aktywności** |
| Godziny z harmonogramu studiów | 60 |
| Inne z udziałem nauczyciela akademickiego  (np. udział w konsultacjach, udział w kolokwium) | 5 |
| Godziny niekontaktowe:  - praca własna studenta  - przeczytanie zalecanej literatury  - przygotowanie i realizacja projektu w postaci prezentacji wraz z dołączonym referatem opisującym zadany problem. | 60 |
| SUMA GODZIN | 125 |
| **SUMARYCZNA LICZBA PUNKTÓW ECTS** | 5 |

*\* Należy uwzględnić, że 1 pkt ECTS odpowiada 25-30 godzin całkowitego nakładu pracy studenta.*

6. PRAKTYKI ZAWODOWE W RAMACH PRZEDMIOTU

|  |  |
| --- | --- |
| wymiar godzinowy | nie dotyczy |
| zasady i formy odbywania praktyk | nie dotyczy |

7. LITERATURA

|  |
| --- |
| Literatura podstawowa:  D. Jaruga, Komunikacja sieciowa źródła informacji big data = Network communication: big data information sources, Wydawnictwo Naukowe i Edukacyjne SBP, Warszawa 2021.  Firlej-Buzon (red.), Big data w humanistyce i naukach społecznych, Wydawnictwo Naukowe i Edukacyjne SBP, Warszawa 2020. |
| Literatura uzupełniająca:  J. Hyman, Microsoft Power BI dla bystrzaków, Helion, Gliwice 2023.  J. Kozłowski, Historia i perspektywy analizy dużych wolumenów i strumieni danych, Wydawnictwo Akademii Sztuki Wojennej, Warszawa 2020.  J. Wątroba, Zastosowania statystyki i data mining w badaniach naukowych, StatSoft Polska, Kraków 2018. |

Akceptacja Kierownika Jednostki lub osoby upoważnionej

1. W przypadku ścieżki kształcenia prowadzącej do uzyskania kwalifikacji nauczycielskich uwzględnić również efekty uczenia się ze standardów kształcenia przygotowującego do wykonywania zawodu nauczyciela. [↑](#footnote-ref-1)