



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Ekstrakcja wiedzy [S2SI1E>EKW]

Przedmiot

Kierunek studiów

Sztuczna inteligencja/Artificial Intelligence

Rok/Semestr

1/2

Studia w zakresie (specjalność)

–

Profil studiów

ogólnoakademicki

Poziom studiów

drugiego stopnia

Język oferowanego przedmiotu

angielski

Forma studiów

stacjonarne

Wymagalność

obligatoryjny

Liczba godzin

Wykład

15

Laboratorium

0

Inne

0

Ćwiczenia

0

Projekty/seminaria

15

Liczba punktów ECTS

3,00

Koordynatorzy

dr Rafał Bachorz

rafal.bachorz@put.poznan.pl

dr hab. inż. Agnieszka Ławrynowicz prof. PP

agnieszka.lawrynowicz@put.poznan.pl

Wykładowcy

Wymagania wstępne

Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z zakresu: technologii webowych (w tym formatów reprezentacji danych XML, JSON), logiki, systemów bazodanowych, sztucznej inteligencji, przetwarzania języka naturalnego oraz programowania w języku Python. Student powinien również rozumieć konieczność poszerzania swoich kompetencji i być skłonny do współpracy w zespole. Ponadto w zakresie kompetencji społecznych student powinien wykazywać takie postawy jak: uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.

Cel przedmiotu

1. Przekazanie studentom wiedzy na temat grafów wiedzy i ekstrakcji wiedzy. 2. Wykształcenie u studentów umiejętności rozwiązywania problemów z zakresu wykorzystania i implementacji metod ekstrakcji wiedzy i systemów intensywnie wykorzystujących wiedzę. 3. Rozwijanie u studentów umiejętności stosowania metod sztucznej inteligencji i analizy danych do ekstrakcji wiedzy i konstruowania grafów wiedzy. 4. Rozwijanie u studentów umiejętności pracy w zespole.

Przedmiotowe efekty uczenia się

Wiedza Student:

1. Ma zaawansowaną i pogłębianą wiedzę w zakresie szeroko rozumianych systemów informacyjnych z wykorzystaniem baz wiedzy, teoretycznych podstaw ich budowy oraz metod, narzędzi i środowisk programistycznych wykorzystywanych do ich realizacji [K2st_W1].
2. Ma uporządkowaną i podbudowaną teoretycznie wiedzę ogólną związaną z kluczowymi zagadnieniami z zakresu pozyskiwania wiedzy [K2st_W2].
3. Ma zaawansowaną wiedzę szczegółową związaną z wybranymi zagadnieniami z zakresu ekstrakcji wiedzy [K2st_W3].
4. Ma wiedzę o trendach rozwojowych i najistotniejszych nowych osiągnięciach w informatyce i innych wybranych pokrewnych dyscyplinach naukowych w zakresie ekstrakcji wiedzy [K2st_W4].
5. Zna zaawansowane metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich i prowadzeniu prac badawczych w zakresie ekstrakcji wiedzy [K2st_W6].

Umiejętności Student:

1. Potrafi pozyskiwać informacje z literatury, baz danych i innych źródeł, integrować je, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski oraz formułować i w pełni uzasadniać opinie [K2st_U1].
2. Potrafi zaplanować i przeprowadzić eksperymenty, w tym pomiary i symulacje komputerowe, interpretować uzyskane wyniki i wyciągać wnioski oraz formułować i weryfikować hipotezy dotyczące złożonych problemów inżynierskich w zakresie pozyskiwania wiedzy oraz prostych problemów badawczych [K2st_U3].
3. Potrafi wykorzystać metody analityczne, symulacyjne i eksperymentalne do formułowania i rozwiązywania zadań inżynierskich w zakresie pozyskiwania wiedzy i prostych problemów badawczych [K2st_U4].
4. Potrafi - przy formułowaniu i rozwiązywaniu zadań z zakresu ekstrakcji wiedzy - integrować wiedzę z różnych obszarów informatyki (a w razie potrzeby także wiedzę z innych dyscyplin naukowych) i stosować podejście systemowe, uwzględniając także aspekty pozatechniczne [K2st_U5].
5. Potrafi ocenić przydatność i możliwość zastosowania nowych osiągnięć (metod i narzędzi) modelowania i inżynierii wiedzy oraz nowych produktów informatycznych wykorzystujących ontologie i grafy wiedzy [K2st_U6].
6. Potrafi krytycznie analizować istniejące rozwiązania techniczne w zakresie pozyskiwania wiedzy i proponować ich ulepszenia [K2st_U8].
7. Potrafi ocenić przydatność metod i narzędzi do rozwiązania zadania inżynierskiego polegającego na budowie lub ocenie systemu informacyjnego lub jego komponentów (w szczególności ontologii, grafów wiedzy lub innych artefaktów z obszaru reprezentacji wiedzy), w tym rozpoznać ograniczenia tych metod i narzędzi [K2st_U9].
8. Potrafi - wykorzystując m.in. nowe koncepcyjnie metody - rozwiązywać złożone zadania informatyczne, w tym zadania niestandardowe oraz zadania z komponentem badawczym, w szczególności w zakresie ekstrakcji wiedzy [K2st_U10].

Kompetencje społeczne Student:

1. Rozumie, że w informatyce wiedza i umiejętności bardzo szybko się dezaktualizują [K2st_K1].
2. Rozumie znaczenie wykorzystania najnowszej wiedzy z zakresu informatyki, zwłaszcza w obszarze ekstrakcji wiedzy, w rozwiązywaniu problemów badawczych i praktycznych [K2st_K2].

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów: na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach i aktywności.
- b) w zakresie projektu: na podstawie oceny bieżącego postępu realizacji zadań cząstkowych.

Ocena podsumowująca:

- a) w zakresie wykładów weryfikacja założonych efektów kształcenia realizowana jest przez egzamin.
- b) w zakresie projektu weryfikacja założonych efektów kształcenia realizowana jest poprzez: ocenę umiejętności związanych z realizacją projektu.

Próg zaliczeniowy: 50% punktów.

Treści programowe

Grafy wiedzy (modele oparte o encje i oparte o własności)
Inżynieria wiedzy (metodyki, narzędzia, dobre praktyki)

Rozpoznawanie encji
 Ekstrakcja relacji
 Wiązanie encji i tagowanie semantyczne
 Ulepszanie grafu wiedzy (przewidywanie powiązań, klasyfikacja encji i trójek, weryfikacja faktów)
 Pozyskiwanie wiedzy i konstruowanie jej z wcześniej wytrenowanych dużych modeli językowych (GPT-3 itp.)

Tematyka zajęć

Grafi wiedzy (modele oparte o encje i oparte o własności)
 Inżynieria wiedzy (metodyki, narzędzia, dobre praktyki)
 Rozpoznawanie encji
 Ekstrakcja relacji
 Wiązanie encji i tagowanie semantyczne
 Ulepszanie grafu wiedzy (przewidywanie powiązań, klasyfikacja encji i trójek, weryfikacja faktów)
 Pozyskiwanie wiedzy i konstruowanie jej z wcześniej wytrenowanych dużych modeli językowych (GPT-3 itp.)

Metody dydaktyczne

Wykład: prezentacja multimedialna, pokaz przykładowych rozwiązań, quizy
 Projekt: czytanie i analiza literatury na temat projektu, projektowanie rozwiązania, realizacja, ocena, praca w zespole

Literatura

Podstawowa 1. Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, Axel-Cyrille Ngonga Ngomo, Axel Polleres, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, Antoine Zimmermann (2021) Knowledge Graphs, Synthesis Lectures on Data, Semantics, and Knowledge, No. 22, 1–237, DOI: 10.2200/S01125ED1V01Y202109DSK022, Morgan & Claypool, available at <https://kgbook.org>
 2. Tom Heath and Christian Bizer, Linked Data: Evolving the Web into a Global Data Space (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011, available at <http://linkeddatatbook.com/book>
 Uzupełniająca 1. Mayank Kejriwal, Craig A. Knoblock, and Pedro Szekely (Eds.). 2021. Knowledge Graphs: Fundamentals, Techniques, and Applications. The MIT Press.
 2. Michael Uschold, Demystifying OWL for the Enterprise, Morgan Claypool Publishers, 2018
 3. Dean Allemang, Jim Hendler, Fabien Gandon, Semantic Web for the Working Ontologist, Third Edition, ACM Books, 2020
 4. Robert Arp, Barry Smith, and Andrew D. Spear. 2015. Building Ontologies with Basic Formal Ontology. The MIT Press.
 5. Agnieszka Ławrynowicz, Semantic data mining. An ontology-based approach, Studies on the Semantic Web, vol. 29. IOS Pres/AKA Verlag 2017
 6. Altinok, Duygu. Mastering spaCy: An end-to-end practical guide to implementing NLP applications using the Python ecosystem. Packt Publishing Ltd, 2021.
 7. Sandra Kublik, Shubham Saboo: GPT-3: Building Innovative NLP Products Using Large Language Models, O'Reilly Media, 2022

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	75	3,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	30	1,50
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	45	1,50