



KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Analiza danych i sieci semantyczne dla Internetu Przedmiotów [S2Inf1-IP>ADISS]

Przedmiot

Kierunek studiów
Informatyka

Rok/Semestr
1/1

Studia w zakresie (specjalność)
Internet Przedmiotów

Profil studiów
ogólnoakademicki

Poziom studiów
drugiego stopnia

Język oferowanego przedmiotu
polski

Forma studiów
stacjonarne

Wymagalność
obligatoryjny

Liczba godzin

Wykład
30

Laboratorium
30

Inne
0

Ćwiczenia
0

Projekty/seminaria
0

Liczba punktów ECTS

5,00

Koordynatorzy

dr inż. Tomasz Łukaszewski
tomasz.lukaszewski@put.poznan.pl

dr inż. Jarosław Bąk
jaroslaw.bak@put.poznan.pl

Wykładowcy

Wymagania wstępne

Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę w zakresie technologii internetowych (w tym języka XML), podstaw logiki i baz danych oraz programowania w językach Python i Java. Powinien również posiadać umiejętność pozyskiwania informacji ze wskazanych źródeł oraz mieć gotowość do podjęcia współpracy w ramach zespołu.

Cel przedmiotu

Przekazanie studentom wiedzy i umiejętności w zakresie analizy danych z wykorzystaniem technik uczenia maszynowego: klasyfikacja, zarządzanie wartościami nieznanymi cech, mapowanie wartości cech, skalowanie wartości cech. Przekazanie studentom podstawowej wiedzy w zakresie technologii semantycznych, w tym koncepcji Web 3.0 (semantyczna sieć WWW) oraz mechanizmów semantycznej integracji danych. Rozwijanie u studentów umiejętności rozwiązywania problemów w zakresie sposobów użytkowania i projektowania systemów wykorzystujących technologie semantyczne.

Przedmiotowe efekty uczenia się

Wiedza:

1. ma zaawansowaną wiedzę szczegółową dotyczącą klasyfikacji, wstępnego przetworzenia danych, selekcji cech, analizy skupień, technologii semantycznych, web 3.0.
2. ma wiedzę o trendach rozwojowych i nowych osiągnięciach w uczeniu maszynowym i technologiach semantycznych.
3. zna zaawansowane metody, techniki i narzędzia stosowane przy rozwiązywaniu złożonych zadań inżynierskich z obszaru informatyki dotyczącego uczenia maszynowego i technologii semantycznych.

Umiejętności:

1. potrafi planować i przeprowadzać eksperymenty w obszarze uczenia maszynowego, interpretować uzyskane wyniki i wyciągać wnioski.
2. potrafi ocenić przydatność metod i narzędzi technologii semantycznych w internecie przedmiotów.

Kompetencje społeczne:

rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe.

Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Wiedza nabyta w ramach wykładu jest weryfikowana na teście zaliczeniowym. Próg zaliczeniowy: 50% punktów. Opcjonalnie możliwość podwyższenia oceny na drodze egzaminu ustnego. Lista zagadnień zostanie podana studentom z odpowiednim wyprzedzeniem czasowym.

Umiejętności nabyte w ramach zajęć laboratoryjnych weryfikowane są na podstawie miniprojektów będących wynikiem analizy wskazanego problemu związanego z uczeniem maszynowym i technologiami semantycznymi.

Treści programowe

Program z zakresu analizy danych obejmuje uczenie maszynowe: klasyfikacja, wstępne przetworzenie danych.

Program z zakresu sieci semantycznych obejmuje semantyczne sieci WWW: modele danych, reprezentacja wiedzy, język zapytań.

Tematyka zajęć

Tematyka wykładu z zakresu analizy danych: wprowadzenie do uczenia maszynowego, problematyka klasyfikacji jako jednego z paradygmatów uczenia maszynowego, klasyfikator najbliższych sąsiadów, drzewa decyzyjne oraz preprocessing danych w zakresie zarządzanie wartościami nieznanymi cech, mapowania wartości cech, skalowania wartości cech, odkrywania cech.

Tematyka wykładu z zakresu sieci semantycznych: pojęcie semantycznej sieci WWW i sieci WWW rzeczy, warstwowa architektura języków semantycznej sieci WWW, trójkowy model danych, język opisu zasobów RDF, reprezentacja wiedzy za pomocą ontologii, odpytywanie heterogenicznych źródeł wiedzy za pomocą języka SPARQL, dostęp do danych poprzez ontologię (R2RML), modelowanie metadanych i inżynierię wiedzy, przykłady aktualnych inicjatyw związanych z semantyczną integracją danych i wiedzy (np. <http://schema.org>), ontologie i schematy metadanych dot. sieci sensorów (W3C SSN) oraz wykorzystanie technologii semantycznych w Internecie Przedmiotów.

Zajęcia laboratoryjne: pogłębienie zagadnień omawianych na wykładach na drodze rozwiązywania praktycznych problemów. W zakresie analizy danych zastosowano m.in. bibliotekę scikit-learn dla języka Python oraz środowisko Jupyter (Jupyter notebooks). W zakresie sieci semantycznych wykorzystano m.in. edytor ontologii (Protégé), modelowanie metadanych (schema.org, JSON-LD), mapowanie OBDA (plugin dla Protégé), bazę danych H2, przetwarzanie danych semantycznych (repozytoria trójek, Jena Fuseki), język SPARQL.

Metody dydaktyczne

Wykład: prezentacja multimedialna

Ćwiczenia laboratoryjne: ćwiczenia praktyczne, dyskusja, praca w zespole

Literatura

Podstawowa

1. Python. Uczenie maszynowe, Wydanie II, Sebastian Raschka, Bahid Mirjalili, Helion 2019
2. Ontologie w systemach informatycznych, Krzysztof Goczyła, EXIT 2011
3. Linked Data: Evolving the Web into a Global Data Space (1st edition). Tom Heath and Christian Bizer, Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011, <http://linkeddatabook.com/book>

Uzupełniająca

1. Naczelny Algorytm. Jak jego odkrycie zmieni nasz świat, Pedro Domingos, Helion 2016
2. Semantic Web for the Working Ontologist, Dean Allemang and Jim Hendler, Morgan Kaufmann 2008
3. Demystifying OWL for the Enterprise, Michael Uschold, Morgan & Claypool Publishers, 2018
4. An Introduction to Ontology Engineering. Keet, C.M. College Publications, volume 20, November 2018
5. Internet rzeczy. Budowa sieci z wykorzystaniem technologii webowych i Raspberry Pi, Dominique Guinard, Vlad Trifa, Helion, 2017
6. Semantic data mining. An ontology-based approach. Agnieszka Ławrynowicz. Studies on the Semantic Web, Vol. 29. IOS Pres/AKA Verlag 2017

Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	125	5,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	60	2,50
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	65	2,50