

KARTA OPISU MODUŁU KSZTAŁCENIA		
Nazwa modułu/przedmiotu Natural language processing		Kod 1010511361010510145
Kierunek studiów Informatyka	Profil kształcenia (ogólnoakademicki, praktyczny) ogólnoakademicki	Rok / Semestr 3 / 6
Ścieżka obieralności/specjalność -	Przedmiot oferowany w języku: polski	Kurs (obligatoryjny/obieralny) obieralny
Stopień studiów: I stopień	Forma studiów (stacjonarna/niestacjonarna) stacjonarna	
Godziny Wykłady: 30 Ćwiczenia: - Laboratoria: 30 Projekty/seminaria: -		Liczba punktów 4
Status przedmiotu w programie studiów (podstawowy, kierunkowy, inny) (ogólnouczelniany, z innego kierunku) kierunkowy z danego kierunku		
Obszar(y) kształcenia i dziedzina(y) nauki i sztuki		Podział ECTS (liczba i %)
Odpowiedzialny za przedmiot / wykładowca: Odpowiedzialny za przedmiot / wykładowca:		
dr hab. inż. Agnieszka Ławrynowicz email: alawrynowicz@cs.put.poznan.pl tel. (+48) 61 6653026 Wydział Informatyki ul. Piotrowo 3, 60-965 Poznań		mgr inż. Dawid Wiśniewski email: dwisniewski@cs.put.poznan.pl tel. (+48) 61 6653053 Wydział Informatyki ul. Piotrowo 3, 60-965 Poznań
Wymagania wstępne w zakresie wiedzy, umiejętności, kompetencji społecznych:		
1	Wiedza:	Student rozpoczynający ten przedmiot powinien posiadać podstawową wiedzę z zakresu programowania (zdobytą na zajęciach z przedmiotów Wprowadzenie do informatyki/Podstawy programowania), podstaw logiki (zdobytą na zajęciach z przedmiotu Logika obliczeniowa), podstaw sztucznej inteligencji (zdobytą na przedmiocie Sztuczna Inteligencja) oraz statystyki i analizy danych (zdobytą na zajęciach z przedmiotu Statystyka i analiza danych).
2	Umiejętności:	Student powinien posiadać umiejętność rozwiązywania podstawowych problemów z zakresu implementacji i oceny kosztu działania prostych algorytmów oraz umiejętność pozyskiwania informacji ze wskazanych źródeł.
3	Kompetencje społeczne	Student powinien również rozumieć konieczność poszerzania swoich kompetencji i wykazywać gotowość do podjęcia współpracy w ramach zespołu. Ponadto w zakresie kompetencji społecznych student musi prezentować takie postawy jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek dla innych ludzi.
Cel przedmiotu:		
<ol style="list-style-type: none"> Przekazanie studentom podstawowej wiedzy na temat przetwarzania języka naturalnego. Rozwijanie u studentów umiejętności rozwiązywania prostych problemów w zakresie sposobów użytkowania i implementacji metod i systemów wykorzystujących przetwarzanie języka naturalnego. Rozwijanie u studentów umiejętności doboru odpowiednich form reprezentacji wiedzy do rozwiązywanego problemu i ich wykorzystania do modelowania i przetwarzania wiedzy. Rozwijanie u studentów umiejętności zastosowania metod sztucznej inteligencji i analizy danych do przetwarzania języka naturalnego. Kształtowanie u studentów umiejętności pracy zespołowej. 		
Efekty kształcenia i odniesienie do kierunkowych efektów kształcenia		
Wiedza:		
<ol style="list-style-type: none"> ma rozszerzoną i pogłębioną wiedzę z matematyki przydatną do formułowania i rozwiązywania złożonych zadań informatycznych dotyczących m.in. programowania w logice, rachunku prawdopodobieństwa, formalnej specyfikacji i weryfikacji oprogramowania - [K1st_W1] ma uporządkowaną i podbudowaną teoretycznie wiedzę ogólną w zakresie kluczowych zagadnień z zakresu przetwarzania języka naturalnego, oraz wiedzę szczegółową z tego zakresu - [K1st_W4] ma wiedzę o istotnych kierunkach rozwoju i najważniejszych osiągnięciach informatyki z zakresu przetwarzania języka naturalnego - [K1st_W5] ma podstawową wiedzę o cyklu życia systemów informatycznych przetwarzania języka naturalnego, zarówno sprzętowych jak i programowych, a w szczególności o zachodzących w nich kluczowych procesach - [K1st_W6] zna podstawowe techniki, metody oraz narzędzia wykorzystywane w procesie rozwiązywania zadań informatycznych z zakresu przetwarzania języka naturalnego - [K1st_W7] 		
Umiejętności:		

<p>1. potrafi pozyskiwać informacje z zakresu przetwarzania języka naturalnego z różnych źródeł, w tym z literatury oraz baz danych, zarówno w języku polskim jak i w języku angielskim, właściwie je integrować, dokonywać ich interpretacji i krytycznej oceny, wyciągać wnioski - [K1st_U1]</p> <p>2. potrafi odpowiednio posługiwać się technikami informacyjno-komunikacyjnymi, znajdującymi zastosowanie na różnych etapach realizacji przedsięwzięć informatycznych z zakresu przetwarzania języka naturalnego - [K1st_U2]</p> <p>3. potrafi, formułując i rozwiązując zadania informatyczne z zakresu przetwarzania języka naturalnego, zastosować odpowiednio dobrane metody, w tym metody analityczne lub eksperymentalne - [K1st_U4]</p> <p>4. potrafi zaprojektować oraz zrealizować system informatyczny z zakresu przetwarzania języka naturalnego, dobierając odpowiedni język programowania oraz używając właściwych metod, technik i narzędzi - [K1st_U10]</p> <p>5. ma umiejętność formułowania algorytmów i ich implementacji z użyciem przynajmniej jednego z popularnych narzędzi - [K1st_U11]</p> <p>6. potrafi planować i realizować proces własnego permanentnego uczenia się z przetwarzania języka naturalnego - [K1st_U19]</p>
<p>Kompetencje społeczne:</p>
<p>1. rozumie, że w informatyce wiedza i umiejętności z zakresu przetwarzania języka naturalnego bardzo szybko stają się przestarzałe - [K1st_K1]</p> <p>2. ma świadomość znaczenia wiedzy w rozwiązywaniu problemów inżynierskich z zakresu przetwarzania języka naturalnego - [K1st_K2]</p> <p>3. potrafi myśleć i działać w sposób przedsiębiorczy, m.in. znajdując komercyjne zastosowania dla tworzonego oprogramowania, mając na uwadze nie tylko korzyści biznesowe, ale również społeczne prowadzonej działalności - [K1st_K3]</p>

<p style="text-align: center;">Sposoby sprawdzenia efektów kształcenia</p>
<p>Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:</p> <p>Ocena formująca:</p> <p>a) w zakresie wykładów:</p> <ul style="list-style-type: none">? na podstawie aktywnego uczestnictwa w wykładach.? na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach (quizz) <p>b) w zakresie laboratoriów:</p> <ul style="list-style-type: none">? na podstawie oceny bieżącego postępu realizacji zadań cząstkowych,? na podstawie postępów projektu. <p>Ocena podsumowująca:</p> <p>a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none">? podsumowanie punktów i omówienie ?quizzów? z wykładu,? prezentacje studentek i studentów, na wybrane tematy związane z przetwarzaniem języka naturalnego,? na ostateczną ocenę w zakresie wykładów składają się: punkty z ?quizzów? dostępnych po każdym wykładzie, punkty za aktywność, punkty za prezentację. <p>b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:</p> <ul style="list-style-type: none">? ocenę umiejętności związanych z realizacją ćwiczeń laboratoryjnych,? ocenę wykonania zadań realizowanych częściowo w trakcie laboratoriów i w części po ich zakończeniu,? ocenę projektu. <p>Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:</p> <ul style="list-style-type: none">? omówienia dodatkowych aspektów zagadnienia,? wykazanie się ciekawymi umiejętnościami ponadprogramowymi,? efektywność zastosowania zdobytej wiedzy podczas rozwiązywania zadanego problemu,? uwagi prowadzące do udoskonalenia materiałów dydaktycznych lub procesu dydaktycznego.
<p style="text-align: center;">Treści programowe</p>
<p>W ramach wykładu przedstawiane są następujące zagadnienia:</p> <ul style="list-style-type: none">? Podstawowa architektura systemów przetwarzania języka naturalnego? Podstawowe techniki przetwarzania tekstu (wyrażenia regularne, filtrowanie wyrazów funkcyjnych, segmentacja, lematyzacja, odległość edycyjna)? Klasyfikacja tekstu? Analiza gramatyczna (rozpoznawanie części mowy, tagowanie POS, parsery zależnościowe)? Statystyczne modelowanie języka, rozproszone reprezentacje: n-gram, word2vec, GloVe, softmax? Reprezentacja wiedzy: sieci semantyczne, grafy wiedzy, tezaury, słowniki (WordNet, FrameNet, DBpedia, schema.org, YAGO, inne)? Ekstrakcja informacji i wiedzy z danych tekstowych (wykrywanie encji nazwanych, ekstrakcja relacji)? Zastosowanie deep learning do przetwarzania języka naturalnego (GRU, LSTM)

?	Analiza sentymentu i opinii	
?	Automatyczne generowanie streszczeń	
?	Odpowiadanie na pytania (question answering)	
?	Systemy dialogowe	
<p>Program laboratorium obejmuje zagadnienia podstawowych metod przetwarzania języka naturalnego z wykorzystaniem popularnych bibliotek i zestawów narzędziowych (biblioteka Pythona NLTK), zagadnienia zastosowania deep learning do przetwarzania języka naturalnego (TensorFlow/Theano/Caffe/Keras) oraz case studies w zakresie zastosowań poznanych metod (analiza sentymentu i opinii, automatyczne generowanie streszczeń, odpowiadanie na pytania):</p>		
?	Modele językowe, tokenizacja, lematyzacja, sentence splitting jako wprowadzenie do przetwarzania języka naturalnego	
?	Szybkie wyszukiwanie w tekście, wyrażenia regularne	
?	Wyszukiwanie przybliżone - odległości edycyjne (Levenshteina), uwzględnianie literówek poprzez wyrażenia regularne	
?	Klasyfikacja tekstu - metoda TF-IDF	
?	Topic modeling - metoda LDA	
?	Analiza gramatyczna - parsery zależnościowe i tagowanie POS	
?	Wykrywanie encji nazwanych, wykrywanie relacji	
?	Analiza sentymentu i opinii	
?	Metody generowania streszczeń	
?	Odpowiadanie na pytania: metody bazujące na wyszukiwaniu informacji (Information Retrieval), metody wykorzystujące słowniki, WordNet oraz metody z wykorzystaniem ekstrakcji relacji, DBpedii, FrameNetu	
?	Problem rzadkich reprezentacji - word embeddings	
?	Wprowadzenie do sieci neuronowych (deep learning) w kontekście przetwarzania tekstu (GRU, LSTM)	
?	Zastosowanie sieci neuronowych (deep learning) do wybranego zadania przetwarzania tekstu	
Literatura podstawowa:		
Literatura uzupełniająca:		
Bilans nakładu pracy przeciętnego studenta		
Czynność	Czas (godz.)	
1. udział w zajęciach laboratoryjnych	30	
2. przygotowanie do ćwiczeń laboratoryjnych	3	
3. dokończenie (w ramach pracy własnej) zadań z ćwiczeń laboratoryjnych	22	
4. udział w konsultacjach związanych z realizacją procesu kształcenia, w szczególności ćwiczeń laboratoryjnych / projektu	2 5	
5. rozwiązanie quizu z wykładu	4	
6. przygotowanie do testu	30	
7. udział w wykładach	4	
8. zapoznanie się ze wskazaną literaturą / materiałami dydaktycznymi (10 stron tekstu naukowego = 1 godz.), 40 stron		
Obciążenie pracą studenta		
forma aktywności	godzin	ECTS
Łączny nakład pracy	100	4
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	62	2
Zajęcia o charakterze praktycznym	55	2