

1. potrafi pozyskiwać informacje 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, oraz wyczerpująco uzasadniać formułowane przez siebie opinie - [K1st_U1]
2. potrafi odpowiednio posługiwać się technikami informacyjno-komunikacyjnymi, znajdującymi zastosowanie na różnych etapach realizacji przedsięwzięć informatycznych - [K1st_U2]
3. potrafi, formułując i rozwiązując zadania informatyczne, zastosować odpowiednio dobrane metody, w tym metody analityczne, symulacyjne lub eksperymentalne - [K1st_U4]
4. potrafi - zgodnie z zadaną specyfikacją - zaprojektować oraz zrealizować urządzenie lub szeroko rozumiany system informatyczny, dobierając język programowania odpowiedni do danego zadania programistycznego oraz używając właściwych metod, technik i narzędzi - [K1st_U10]
5. ma umiejętność formułowania algorytmów i ich implementacji z użyciem przynajmniej jednego z popularnych narzędzi - [K1st_U11]
6. potrafi planować i realizować proces własnego permanentnego uczenia się oraz zna możliwości dalszego dokształcania się (studia II i III stopnia, studia podyplomowe, kursy i egzaminy przeprowadzane przez uczelnie, firmy i organizacje zawodowe) - [K1st_U19]

Kompetencje społeczne:

1. rozumie, że w informatyce wiedza i umiejętności bardzo szybko stają się przestarzałe - [K1st_K1]
2. ma świadomość znaczenia wiedzy w rozwiązywaniu problemów inżynierskich oraz zna przykłady i rozumie przyczyny wadliwie działających systemów informatycznych, które doprowadziły do poważnych strat finansowych, społecznych lub też do poważnej utraty zdrowia, a nawet życia - [K1st_K2]

Sposoby sprawdzenia efektów kształcenia

Efekty kształcenia przedstawione wyżej weryfikowane są w następujący sposób:

Ocena formująca:

- a) w zakresie wykładów:
 - ? na podstawie aktywnego uczestnictwa w wykładach.
 - ? na podstawie odpowiedzi na pytania dotyczące materiału omówionego na poprzednich wykładach ("quizy").
- b) w zakresie laboratoriów:
 - ? na podstawie aktywnego uczestnictwa w laboratoriach i oceny bieżącego postępu realizacji zadań cząstkowych.

Ocena podsumowująca:

- a) w zakresie wykładów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ? ocenę wiedzy na podstawie prezentacji wybranego zagadnienia,
 - ? podsumowanie punktów i omówienie "quizów" z wykładu,
 - ? na ostateczną ocenę w zakresie wykładów składają się: punkty zdobyte na podstawie aktywnego uczestnictwa w wykładzie, punkty z "quizów" dostępnych po wybranych wykładach, punkty z prezentacji.
- b) w zakresie laboratoriów weryfikowanie założonych efektów kształcenia realizowane jest przez:
 - ? ocenę umiejętności związanych z aktywną realizacją ćwiczeń laboratoryjnych,
 - ? ocenę wykonania zadań realizowanych częściowo w trakcie laboratoriów i w części po ich zakończeniu,
 - ? ocenę wykonania i prezentacji projektu realizowanego częściowo w trakcie laboratoriów i w części po ich zakończeniu,

Uzyskiwanie punktów dodatkowych za aktywność podczas zajęć, a szczególnie za:

- ? omówienia dodatkowych aspektów zagadnienia,
- ? wykazanie się ciekawymi umiejętnościami ponadprogramowymi,
- ? uwagi prowadzące do udoskonalenia materiałów dydaktycznych lub procesu dydaktycznego.

Treści programowe

W ramach wykładu przedstawiane są następujące zagadnienia:

- ? 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

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):

- ? 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:

1. Natural Language Processing with Python, Steven Bird, Ewan Klein, and Edward Loper, O'Reilly Media, 2009, dostępna online <http://www.nltk.org/book/>
2. Speech and Language Processing (3rd ed. draft), Dan Jurafsky and James H. Martin. Draft chapters in progress, August 28, 2017 (dostępna online)

Literatura uzupełniająca:

1. Foundations of Statistical Natural Language Processing, Chris Manning and Hinrich Schütze, MIT Press. Cambridge, MA: May 1999, <http://nlp.stanford.edu/fsnlp/>
2. The Text Mining Handbook, Ronen Feldman, James Sanger, Cambridge University Press, 2007
3. Inżynieria lingwistyczna. Komputerowe przetwarzanie tekstów w języku naturalnym, Agnieszka Mykowiecka, 2007, Wydawnictwo PJWSTK, Seria: Podręczniki akademickie.
4. Deep Learning, Ian Goodfellow, Yoshua Bengio, Aaron Courville, MIT Press 2016, <http://www.deeplearningbook.org>
5. Semantic Web for the Working Ontologist, Dean Allemang and Jim Hendler, Morgan Kaufmann, 2008
6. 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, 2011, 1:1, 1-136. Morgan & Claypool, dostępna online <http://linkeddatadatabook.com>
7. Voice Application Development for Android, Michael F. McTear, Zoraida Callejas, PACKT Publishing, 2013

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
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	62
Zajęcia o charakterze praktycznym	52