



## KARTA OPISU PRZEDMIOTU - SYLABUS

Nazwa przedmiotu

Przetwarzanie języka naturalnego [S1S1E>PJN]

### Przedmiot

Kierunek studiów

Sztuczna inteligencja/Artificial Intelligence

Rok/Semestr

3/6

Studia w zakresie (specjalność)

–

Profil studiów

ogólnoakademicki

Poziom studiów

pierwszego stopnia

Język oferowanego przedmiotu

angielski

Forma studiów

stacjonarne

Wymagalność

obligatoryjny

### Liczba godzin

Wykład

30

Laboratorium

30

Inne

0

Ćwiczenia

0

Projekty/seminaria

0

### Liczba punktów ECTS

4,00

### Koordynatorzy

dr inż. Dawid Wiśniewski

dawid.wisniewski@put.poznan.pl

dr hab. inż. Mikołaj Morzy prof. PP

mikolaj.morzy@put.poznan.pl

dr hab. inż. Agnieszka Ławrynowicz prof. PP

agnieszka.lawrynowicz@put.poznan.pl

### Wykładowcy

### Wymagania wstępne

Student rozpoczynający ten kurs powinien posiadać podstawową wiedzę z zakresu programowania, podstaw logiki, podstawy sztucznej inteligencji oraz statystyki i analizy danych. Uczeń powinien posiadać umiejętność rozwiązywania podstawowych problemów z zakresu implementacji i oceny kosztów prostych algorytmów, oraz umiejętność pozyskiwania informacji ze wskazanych źródeł. Uczeń powinien także rozumieć konieczność poszerzania swoich kompetencji i chęć współpracy w zespole. Co więcej, pod względem kompetencji społecznych uczeń powinien wykazywać się takimi postawami, jak uczciwość, odpowiedzialność, wytrwałość, ciekawość poznawcza, kreatywność, kultura osobista, szacunek do drugiego człowieka.

## Cel przedmiotu

1. Przekazanie studentom podstawowej wiedzy z zakresu przetwarzania języka naturalnego. 2. Wykształcenie u studentów umiejętności rozwiązywania prostych problemów w zakresie stosowania i wdrażania metod oraz systemy wykorzystujące przetwarzanie języka naturalnego. 3. Rozwijanie umiejętności studentów w zakresie stosowania sztucznej inteligencji i metod analizy danych w środowisku naturalnym przetwarzanie języka. 4. Rozwijanie u uczniów umiejętności pracy zespołowej.

## Przedmiotowe efekty uczenia się

Wiedza:

1. Student posiada ustrukturyzowaną i ugruntowaną teoretycznie wiedzę ogólną dotyczącą kluczowych zagadnień nt przetwarzanie języka naturalnego i szczegółową wiedzę na wybrane tematy z tej dziedziny nauki.
2. Student zna główne kierunki rozwoju i najważniejsze osiągnięcia w dziedzinie przetwarzania języka naturalnego i innych dziedzin pokrewnych, takich jak wykorzystanie Metody uczenia maszynowego w przetwarzaniu tekstu i inne zagadnienia sztucznej inteligencji związane z naturą przetwarzania języka.
3. Student ma podstawową wiedzę na temat najważniejszych i najbardziej istotnych kierunków rozwoju zrozumiano ważne osiągnięcia sztucznej inteligencji (w szczególności przetwarzanie języka naturalnego) jako ważna dziedzina informatyki czerpiąca z osiągnięć innych dyscyplin naukowych i dostarcza im rozwiązań o praktycznym potencjale.

Umiejętności:

1. Student potrafi pozyskiwać informacje z różnych źródeł, w tym z literatury i baz danych, zarówno w języku polskim, jak i angielskim, właściwie je integrować, interpretować i krytycznie oceniać, rysować wyciągać wnioski i dokładnie uzasadniać formułowane przez siebie opinie.
2. Student potrafi wykorzystać technologie informacyjno-komunikacyjne na różnych etapach życia realizację projektów informatycznych.
3. Student potrafi formułować i rozwiązywać zadania sztucznej inteligencji przy użyciu odpowiednich metod, w tym metodami analitycznymi, symulacyjnymi lub eksperymentalnymi.
4. Student potrafi zaprojektować i wdrożyć urządzenie lub szeroko pojęty system informatyczny zgodnie z zadaną specyfikacją, wybierając odpowiedni dla zadanego język programowania programowania i stosowania odpowiednich metod, technik i narzędzi sztucznej inteligencji.
5. Student potrafi formułować algorytmy i je implementować z wykorzystaniem co najmniej jednego popularnego narzędzia.
6. Student potrafi pozyskiwać, analizować i przetwarzać zbiory danych tekstowych, zabezpieczać je przed niepowołanymi osobami dostęp do nich i syntezę ich w wiedzę i wnioski przydatne w rozwiązywaniu szerokiego zakresu problemów powstające w pracy informatyków, specjalistów z zakresu sztucznej inteligencji, m.in problemy specyfiki przemysłu, biznesu i administracji
7. Student potrafi wykorzystywać i adaptować modele inteligentnych zachowań oraz narzędzia informatyczne tam, gdzie jest to naturalne wykorzystywane jest przetwarzanie językowe, które symuluje te zachowania
8. Student potrafi posługiwać się technikami informacyjnymi i komunikacyjnymi, zwłaszcza wykorzystującymi naturalne komponentów i narzędzi do przetwarzania języka na różnych etapach projektów informatycznych, w tym przygotowania dobrze udokumentowane badanie problemu, ustna prezentacja, komunikacja ze specjalistą terminologii i omawiania różnych opinii i stanowisk, także w środowisku niespecjalistycznym

Kompetencje społeczne:

1. Student rozumie, że w sztucznej inteligencji wiedza i umiejętności bardzo się dezaktualizują szybko, dostrzegając potrzebę ciągłego dokształcania się i podnoszenia własnych kompetencji.
2. Student ma świadomość znaczenia wiedzy i badań naukowych z nią związanych informatyka i sztuczna inteligencja (w szczególności przetwarzanie języka naturalnego) w rozwiązywaniu zadań praktycznych problemy kluczowe dla funkcjonowania jednostek, firm, organizacji i całości

społeczeństwo

3. Student potrafi myśleć i działać w sposób przedsiębiorczy, tj. znajdować zastosowania komercyjne stworzył systemy sztucznej inteligencji (polegającej na przetwarzaniu języka naturalnego), biorąc pod uwagę nie tylko korzyści ekonomiczne, ale także aspekty prawne i społeczne

### Metody weryfikacji efektów uczenia się i kryteria oceny

Efekty uczenia się przedstawione wyżej weryfikowane są w następujący sposób:

Przedstawione powyżej efekty uczenia się weryfikowane są w następujący sposób:

Ocenianie kształtujące:

a) w zakresie wykładów: na podstawie odpowiedzi na pytania dotyczące materiału omówionego powyżej wykłady (w formie testów) i ćwiczenia.

b) dla laboratoriów: na podstawie oceny bieżącego stanu realizacji zadań cząstkowych.

Ocena podsumowująca:

a) w zakresie zajęć weryfikacja zakładanych efektów kształcenia następuje poprzez: ocenę wiedza i umiejętności zaprezentowane w prezentacji wynikające z realizacji projektu związanego z realizacją projektu

do przetwarzania języka naturalnego oraz podsumowania punktów i omówienia testów z wykładu. Finał na ocenę z wykładów składają się: punkty z testów dostępnych po wybranych wykładach, punkty z prezentacji i punkty za aktywność.

b) w zakresie laboratorium weryfikacja założonych efektów uczenia się odbywa się poprzez: ocenę umiejętności związanych z realizacją ćwiczeń laboratoryjnych, oceną wykonania zadań realizowany częściowo w trakcie laboratoriów, a częściowo po ich zakończeniu, ewaluacja projektu które uczniowie będą realizować, podsumowując zdobytą wiedzę i umiejętności.

Zdobycie dodatkowych punktów za aktywność na zajęciach, w szczególności za: omówienie dodatkowych aspekty problemu, wykazanie ciekawych umiejętności wykraczających poza program, skuteczność zastosowanie zdobytej wiedzy przy rozwiązywaniu problemu.

Ocena zaliczona: 50% punktów.

### Treści programowe

1) Podstawy lingwistyki komputerowej

wprowadzenie do NLP, ogólne omówienie tematu

morfologia, składnia, podział zdań, części mowy i części zdania, tagowanie POS, analizowanie zależności

rzadkie reprezentacje wektorowe (BOW, BONG, TFIDF)

modelowanie języka za pomocą łańcuchów Markowa, CRF

2) Gęste reprezentacje wektorowe

koder-dekoder i model word2vec

osadzanie statyczne: GloVe, FastText, Pointcare, oparte na zależnościach

3) Ekstrakcja informacji (NER)

4) Modele językowe

pierwsze podejścia: RNN, LSTM, GRU

Architektura transformatorowa

BERTologia: BERT, CamemBERT, ROBERT

modele zaawansowane: sieci syjamskie, transformatory zdań, longformery

5) Aplikacje

adnotacja tekstowa dla NLP

zbiory danych, benchmarki, testy

doskonalenie i transfer uczenia się w NLP

### Tematyka zajęć

1) Podstawy lingwistyki komputerowej

wprowadzenie do NLP, ogólne omówienie tematu

morfologia, składnia, podział zdań, części mowy i części zdania, tagowanie POS, analizowanie zależności

rzadkie reprezentacje wektorowe (BOW, BONG, TFIDF)

modelowanie języka za pomocą łańcuchów Markowa, CRF

- 2) Gęste reprezentacje wektorowe  
koder-dekoder i model word2vec  
osadzanie statyczne: GloVe, FastText, Pointcare, oparte na zależnościach
- 3) Ekstrakcja informacji (NER)
- 4) Modele językowe  
pierwsze podejścia: RNN, LSTM, GRU  
Architektura transformatorowa  
BERTologia: BERT, CamemBERT, ROBERT  
modele zaawansowane: sieci syjamskie, transformatory zdań, longformery
- 5) Aplikacje  
adnotacja tekstowa dla NLP  
zbiory danych, benchmarki, testy  
doskonalenie i transfer uczenia się w NLP

## Metody dydaktyczne

wykład: prezentacja multimedialna, demonstracja przykładowych rozwiązań, quizy  
zajęcia laboratoryjne: ćwiczenia praktyczne, dyskusja, praca zespołowa, analiza materiałów multimedialnych

## Literatura

### Podstawowa

1. Speech and Language Processing (3rd ed. draft), Dan Jurafsky and James H. Martin. Draft chapters in progress, Dec 29, 2021, <https://web.stanford.edu/~jurafsky/slp3/>
2. Natural Language Processing with Python, Steven Bird, Ewan Klein, and Edward Loper, O'Reilly Media, 2009, <http://www.nltk.org/book/>

### Uzupełniająca

1. Natural Language Processing in Action. Understanding, analyzing, and generating text with Python, Hobson Lane, Cole Howard, Hannes Hapke, Manning Publications, 2019
2. Practical Natural Language Processing: A Pragmatic Approach to Processing and Analyzing Language Data: A Comprehensive Guide to Building Real-World NLP Systems, Sowmya Vajjala, Bodhisattwa Majumder i in., O'Reilly Media, 2020
3. Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning 1st Edition, Delip Rao, Brian McMahan, O'Reilly Media, 2019
4. Foundations of Statistical Natural Language Processing, Chris Manning and Hinrich Schütze, MIT Press. Cambridge, MA: May 1999, <http://nlp.stanford.edu/fsnlp/>
5. The Text Mining Handbook, Ronen Feldman, James Sanger, Cambridge University Press, 2007
6. Rothman, Denis. Transformers for Natural Language Processing: Build innovative deep neural network architectures for NLP with Python, PyTorch, TensorFlow, BERT, RoBERTa, and more. Packt Publishing Ltd, 2021.
7. Ravichandiran, Sudharsan. Getting Started with Google BERT: Build and train state-of-the-art natural language processing models using BERT. Packt Publishing Ltd, 2021.
8. Altinok, Duygu. Mastering spaCy: An end-to-end practical guide to implementing NLP applications using the Python ecosystem. Packt Publishing Ltd, 2021.

## Bilans nakładu pracy przeciętnego studenta

	Godzin	ECTS
Łączny nakład pracy	100	4,00
Zajęcia wymagające bezpośredniego kontaktu z nauczycielem	62	2,50
Praca własna studenta (studia literaturowe, przygotowanie do zajęć laboratoryjnych/ćwiczeń, przygotowanie do kolokwium/egzaminu, wykonanie projektu)	38	1,50