

Zgłoszenie tematu pracy dyplomowej :: **STUDIA II STOPNIA** ::

na rok akademicki 2023/24

Promotor:	dr hab. Jozef Kapusta, prof. UP
Temat pracy magisterskiej (j. polski oraz j. angielski):	Machine Learning in the Part-Of-Speech Tagging of Texts <i>Wykorzystanie uczenia maszynowego w tagowaniu części mowy</i>
Zakres i oczekiwane rezultaty pracy:	<p>Nowadays, rapidly developing fields of machine learning and natural language processing brings new opportunities for automating text analysis. This thesis will focus on utilizing deep learning for the automatic morphological annotation of texts. Deep learning, as one of the most significant approaches in machine learning, allows for creating more precise and efficient models for various language processing tasks, including morphological and syntactic analysis of texts.</p> <p>The aim of the thesis is to explore existing approaches to the morphological annotation of texts. Within the scope of the work, student will analyze existing morphological taggers for the Polish language, their architecture, performance, and accuracy. Ultimately, all approaches will be evaluated, and the best architecture for Polish language processing will be presented.</p>
*Aspekt naukowy, problemowy pracy:	definition and implementation of a part of speech tagging, implementation natural language processing methods, implementation machine learning methods.
Literatura	<ul style="list-style-type: none"> • Bird, S., Klein E., and Loper, E. (2009). Natural Language Processing with Python - Analyzing Text with the Natural Language Toolkit. O'Reilly Media. • Bengfort, B., Ojeda, T., Bilbro, R. (2018). Applied Text Analysis with Python: Enabling Language - Aware Data Products with Machine Learning, O'Reilly Media, 332 p. • Natural Language Toolkit, online: https://www.nltk.org/ • Kapusta, J., Drlik, M., Munk, M. (2021). Using of N-grams From Morphological Tags for Fake News Classification. PEERJ COMPUTER SCIENCE, PeerJ, 7, e624. • Nagy, K., Kapusta, J. (2021). Improving fake news classification using dependency grammar. PLOS ONE, Public Library of Science, 16(9), e0256940 • Chiche, A., Yitagesu, B. (2022). Part of speech tagging: a systematic review of deep learning and machine learning approaches. JOURNAL OF BIG DATA, 9/1, nu. 10
, **Oprogramowanie, język programowania, środowisko systemowe:	Jupyter Notebook Environment (Python)
**Środowisko uruchomieniowe:	

Zgłoszenie tematu pracy dyplomowej :: **STUDIA II STOPNIA** ::

na rok akademicki 2023/24

Dodatkowe wymagania i uwagi:	English language
------------------------------	------------------

UWAGA:

W polu literatura należy wskazać minimum 1 publikację z listy czasopism punktowanych wg wykazu MEiN z dnia 21 grudnia 2021 r związaną z proponowanym tematem pracy dyplomowej.

* Regulamin studiów § 36 2. Praca dyplomowa na profilu praktycznym, podobnie jak praca inżynierska, powinna mieć charakter aplikacyjny, badawczy, projektowy lub oceniający praktykę w świetle teorii.

** pola opcjonalne