

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):	Large Language Models for Documents Analysis Duże modele językowe do analizy dokumentów
Zakres i oczekiwane rezultaty pracy:	<p>Natural Language Generation systems can produce text for different applications, for example, the generation of sentences in a certain style, machine translation which is nothing but the generation of text in a different language, auto-completion of a sentence given part of that sentence as input, generation of textual summaries, automated chatbots, etc.</p> <p>In the theoretical part: The thesis aim is a summarization of large language models and using these models for language generation.</p> <p>In the practical part: In the practical part of the thesis, the student creates a case study in Python. The case study will focus on using Large language models to analyze documents and answer specific questions regarding that documents.</p>
*Aspekt naukowy, problemowy pracy:	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 • Cagliero, L., La Quatra, M. (2021). Inferring Multilingual Domain-Specific Word Embeddings From Large Document Corpora. IEEE ACCESS, 9, pp. 137309-137321
, **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