

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

na rok akademicki 2024/25

Promotor:	dr hab. Jozef Kapusta, prof. UKEN
Temat pracy magisterskiej (j. polski oraz j. angielski):	<p>Improving Text Classification Based on the Coreference Resolution Method</p> <p>Udoskonalanie klasyfikacji tekstu w oparciu o metodę Coreference Resolution</p>
Zakres i oczekiwane rezultaty pracy:	<p>Given the massive amount of text content, classification methods based on machine learning algorithms is a practical natural language processing problem. For the natural language processing area is potentially useful the coreference resolution. It is the task of clustering mentions in text that refer to the same underlying real-world entities. Coreference resolution aims at resolving repeated references to an object in a document and forms a core component of natural language processing (NLP) research.</p> <p>The aim of the thesis is to determine whether coreference resolution can improve performance measures in classification models.</p> <p>The theoretical part of the thesis will summarise the methods for coreference resolution. New experimental methods from current scientific sources are expected to be explored.</p> <p>The practical part of the thesis will focus on creating two text classifiers. The dataset for the first classifier will contain the text data. The dataset for the second will be the same but with coreference resolution. A student has to apply the selected method for coreference resolution. A student evaluates the results from two classifiers due to performance measures of the classifiers (accuracy, precision, recall, f1-score). The aim of the practical part is to determine the importance of the application of the coreference resolution method on text classification.</p>
*Aspekt naukowy, problemowy pracy:	definition and implementation of the coreference resolution method, implementation methods natural language, 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/ • Ogrodniczuk, M. (2019). Nominal Coreference Resolution for Polish. In. Poznan Studies in Contemporary Linguistics, Vol. 55 / 2, p. 367-396 • Yash, J. (2022). Coreference Resolution [NLP, Python], online: https://blog.devgenius.io/coreference-resolution-nlp-python-

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

na rok akademicki 2024/25

	584c2ec50f5d <ul style="list-style-type: none">Lata, K., Singh, P., Dutta, K. (2022). Mention detection in coreference resolution: survey. In. Applied Intelligence. (Early Access, jan 2022)
, **Oprogramowanie, język programowania, środowisko systemowe:	Jupyter Notebook Environment (Python)
**Środowisko uruchomieniowe:	
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