

## I. Personal and study details

Student's name: **Zviazdou Yauheni** Personal ID number: **507333**  
Faculty / Institute: **Faculty of Electrical Engineering**  
Department / Institute: **Department of Cybernetics**  
Study program: **Open Informatics**  
Specialisation: **Artificial Intelligence and Computer Science**

## II. Bachelor's thesis details

Bachelor's thesis title in English:

**From FastText to Transformer Models, and their Application in Retrieval-Augmented Generation**

Bachelor's thesis title in Czech:

**Od FastText k Transformer model m a jejich aplikace v Retrieval-Augmented generování**

Guidelines:

Review the representation of words, sentences, and paragraphs, progressing from traditional methods like FastText to advanced transformer-based models such as BERT. The primary focus is to evaluate selected representations using analogy tests and confusion matrix. Use the UPV corpus set for evaluation.

In the second part of the study, emphasis will shift towards selecting optimal representations for Retrieval-Augmented Generation (RAG) algorithms. The investigation will determine the most efficient embeddings and optimal text chunk size for question-answering tasks, particularly in the context of natural language answers generation from technical manuals. Conduct a comprehensive evaluation with a particular focus on suggesting an optimal representation model that balances factuality and CPU requirements.

Bibliography / sources:

- [1] FastText documentation and tutorials, <https://fasttext.cc/>
- [2] LangChain documentation, <https://js.langchain.com/docs>
- [3] Word2Vec tutorials, <https://www.tensorflow.org/text/tutorials/word2vec>

Name and workplace of bachelor's thesis supervisor:

**Ing. Jan Šedivý, CSc. Big Data and Cloud Computing CIIRC**

Name and workplace of second bachelor's thesis supervisor or consultant:

Date of bachelor's thesis assignment: **02.01.2024** Deadline for bachelor thesis submission: **24.05.2024**

Assignment valid until: **21.09.2025**

Ing. Jan Šedivý, CSc.  
Supervisor's signature

prof. Dr. Ing. Jan Kybic  
Head of department's signature

prof. Mgr. Petr Páta, Ph.D.  
Dean's signature

## III. Assignment receipt

The student acknowledges that the bachelor's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor's thesis, the author must state the names of consultants and include a list of references.

\_\_\_\_\_  
Date of assignment receipt

\_\_\_\_\_  
Student's signature