Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Adam Zahradník
Study programme:	Applied Computer Science (Single degree study, bachelor I. deg., full time form)
Field of Study:	Computer Science
Type of Thesis:	Bachelor´s thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Exploring Advanced Reasoning Abilities of Large Language Models in Slovak

Annotation: In the recent past, Large Language Models (LLMs) have shown impressive capabilities in tackling various quantitative reasoning and knowledge challenges in fields like mathematics, physics, and computer science. These achievements are typically assessed using benchmarks designed for primary and secondary school levels, often derived from standardized tests. However, this approach poses a problem: standardized tests frequently find their way into the training data of LLMs, leading to skewed performance evaluations. Additionally, the evaluation of LLMs' reasoning abilities primarily occurs in English, raising concerns about the applicability and generalizability of the results.

Aim: The goals of the bachelor's thesis include (but are not limited to)
- analysis of the current state-of-the-art in reasoning capability evaluation
- creation and/or collection of reasoning evaluation datasets in Slovak language based on tasks/exercises in Olympiads in Mathematics, Physics or Informations, as well as various correspondence seminars in the same areas
- evaluation of state-of-the-art LLMs on the prepared datasets
- analysis of the outputs (and error modes) produced by the best performing models

Literature:
Hendrycks, Dan, et al. "Measuring mathematical problem solving with the math dataset." arXiv preprint arXiv:2103.03874 (2021). (https://arxiv.org/pdf/2103.03874.pdf)
Cobbe, Karl, et al. "Training verifiers to solve math word problems." arXiv preprint arXiv:2110.14168 (2021). (https://arxiv.org/pdf/2110.14168.pdf)
Wei, Jason, et al. "Chain-of-thought prompting elicits reasoning in large language models." Advances in Neural Information Processing Systems 35 (2022): 24824-24837. (https://arxiv.org/abs/2201.11903)
Sawada, Tomohiro, et al. "Arb: Advanced reasoning benchmark for large language models." arXiv preprint arXiv:2307.13692 (2023). (https://arxiv.org/pdf/2307.13692.pdf)

Supervisor:	Mgr. Marek Šuppa
Department:	FMFI.KAI - Department of Applied Informatics
Head of department:	doc. RNDr. Tatiana Jajcayová, PhD.

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

Assigned: 15.10.2023

Approved: 16.10.2023 doc. RNDr. Damas Gruska, PhD.
Guarantor of Study Programme

.....
Student

.....
Supervisor