

## Velké jazykové modely v optimalizaci struktury modelů strojového a hlubokého učení

**Školitel:** prof. Ing. Komínková Oplatková Zuzana, Ph.D.

**Konzultant:** Ing. Turečková Alžběta, Ph.D., ---

**Ústav fakulty:** Ústav informatiky a umělé inteligence

**Studijní program:** Informační technologie

### **Anotace:**

Cílem tohoto výzkumu je pokročit v používání velkých jazykových modelů (LLM) při strukturální optimalizaci modelů strojového učení a hlubokého učení. Ústředním bodem tohoto přístupu je integrace poznatků založených na LLM do procesu návrhu modelu, se zvláštním zaměřením na zkoumání schopností LLM identifikovat a navrhnout inovativní vzory struktur. Cílem tohoto zkoumání je výzva prozkoumat a případně rozšířit hranice konvenčního návrhu modelu, včetně vývoje nových struktur a funkčních bloků, jakož i optimalizace hyperparametrů, přizpůsobených specifickým charakteristikám datových sad a chování výsledného modelu. Očekávaným výsledkem je účinnější metodika, která se bude vyznačovat výrazným zvýšením efektivnosti, přesnosti a přizpůsobivosti modelu v různých oblastech a zároveň snížením spotřeby energie při trénování. V rámci výzkumu bude tato metodika navržena a ověřena prostřednictvím komplexních experimentů. Očekávané výsledky by měly obohatit vědeckou komunitu o nové přístupy, metodiky a nabídnout cenné poznatky a praktické podněty pro aplikace LLM v této oblasti.

### **Literatura:**

- [1] KHAN, Wahab, et al. Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends. *Natural Language Processing Journal*, 2023, 100026.
- [2] BU, Kun; LIU, Yuanchao; JU, Xiaolong. Efficient Utilization of Pre-trained Models: A Review of Sentiment Analysis via Prompt Learning. *Knowledge-Based Systems*, 2023, 111148.
- [3] ESPEJEL, Jessica López, et al. GPT-3.5, GPT-4, or BARD? Evaluating LLMs reasoning ability in zero-shot setting and performance boosting through prompts. *Natural Language Processing Journal*, 2023, 5: 100032.
- [4] YANG, Zhen, et al. Improving domain-specific neural code generation with few-shot meta-learning. *Information and Software Technology*, 2023, 107365.
- [5] HE, Xin; ZHAO, Kaiyong; CHU, Xiaowen. AutoML: A survey of the state-of-the-art. *Knowledge-Based Systems*, 2021, 212: 106622.
- [6] NIKITIN, Nikolay O., et al. Automated evolutionary approach for the design of composite machine learning pipelines. *Future Generation Computer Systems*, 2022, 127: 109-125.