

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta aplikovaných věd
Akademický rok: 2021/2022

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Jan PAŠEK**
Osobní číslo: **A20N0098P**
Studijní program: **N3902 Inženýrská informatika**
Studijní obor: **Softwarové inženýrství**
Téma práce: **Generování zdrojových kódů na základě popisu v přirozeném jazyce**
Zadávací katedra: **Katedra informatiky a výpočetní techniky**

Zásady pro vypracování

1. Prozkoumejte stávající metody generování zdrojového kódu algoritmy strojového učení.
2. Na základě studia existujících metod zvolte či navrhnete vhodnou techniku pro generování kódu na základě popisu v přirozeném jazyce. Při volbě techniky se zaměřte na využití moderních předtrénovaných generativních modelů neuronových sítí.
3. Vyberte a získejte datovou sadu zdrojových kódů, která svou velikostí postačuje k předtrénování hluboké neuronové sítě. Pomocí těchto dat předtrénujte vybraný generativní model zdrojového kódu na dostatečně výkonném uzlu gridové infrastruktury MetaCentrum.
4. Implementujte algoritmus pro generování zdrojového kódu z popisu problému v anglickém jazyce. Vyberte a získejte dodatečná data obsahující útržky kódu s odpovídajícím popisem. Získaná data použijte pro dotrénování navržené architektury neuronové sítě.
5. Porovnejte své řešení s již existujícími metodami a proveďte kritické zhodnocení dosažených výsledků.

Rozsah diplomové práce: **doporuč. 50 s. původního textu**
Rozsah grafických prací: **dle potřeby**
Forma zpracování diplomové práce: **tištěná/elektronická**
Jazyk zpracování: **Angličtina**

Seznam doporučené literatury:

dodá vedoucí diplomové práce

Vedoucí diplomové práce: **Ing. Miloslav Konopík, Ph.D.**
Katedra informatiky a výpočetní techniky

Datum zadání diplomové práce: **10. září 2021**
Termín odevzdání diplomové práce: **19. května 2022**

L.S.

Doc. Ing. Miloš Železný, Ph.D.
děkan

Doc. Ing. Přemysl Brada, MSc., Ph.D.
vedoucí katedry

V Plzni dne 11. října 2021