

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Student: Bc. Jan Machynka

Oponent: prof. RNDr. PaedDr. Eva Volná,
PhD.

Studijní program: Informační technologie
Specializace: Softwarové inženýrství
Akademický rok: 2024/2025

Téma diplomové práce: Vizualizace mechanismů pozornosti ve Vision Transformerech

Hodnocení práce:

Úplnost vypracování, aktuálnost a obtížnost řešeného úkolu

Téma diplomové práce je aktuální. Téma interpretovatelnosti složitých modelů umělé inteligence, zejména v oblasti počítačového vidění, je v rychle se vyvíjejícím oboru AI klíčovou výzvou. Nástroje pro vizualizaci mechanismů pozornosti jsou stále ještě ve velmi rané fázi vývoje. Práce reflektuje nejnovější výzkum a publikace v dané oblasti. Tato diplomová práce se zaměřuje na vizualizaci a interpretovatelnost mechanismů pozornosti v architektuře Vision Transformer (ViT).

Způsob a úroveň pojetí řešeného úkolu

Cílem diplomové práce bylo prozkoumat mechanismy pozornosti a jejich interpretovatelnost jak z teoretického hlediska, tak prostřednictvím praktické aplikace vizualizačního nástroje. Předložená diplomová práce představuje kvalitní a komplexní studii v oblasti vizualizace a interpretovatelnosti mechanismů pozornosti ve Vision Transformerech. Diplomant prokázal teoretické znalosti i potřebné praktické dovednosti při vývoji a adaptaci vybraných vizualizačních metod na moderní architektury neuronových sítí. Zároveň systematicky porovnává vizualizační schopnosti jednotlivých modelů a hodnotí jejich efektivitu dle zvolených metrik. Největším přínosem je úspěšná implementace metody LRP pro hierarchický model Hiera, který byl následně podrobně testován.

Úroveň zpracování tématu, přínos diplomanta

Teoretická část představuje základní architekturu ViT a zkoumá jejich relevantní varianty, přičemž systematicky rozebírá dosavadní přístupy k vizualizaci a interpretaci mechanismů pozornosti, a to včetně důkladného zasazení do současného poznání vizualizací a interpretovatelnosti strojového učení. Praktická část popisuje vývoj a implementaci nástroje pro vizualizaci variant ViT, včetně navržené adaptace metody LRP pro hierarchický transformer, a hodnotí klíčové aspekty a omezení použitých metod. Výsledky práce přispívají k hlubšímu pochopení vnitřního fungování modelů a nastiňují směr pro další výzkum v oblasti vizualizací mechanismů pozornosti a jejich interpretovatelnosti.

Formální náležitosti práce, chyby a omyly v technické zprávě

Výběr literatury je dostatečný a odpovídá tématu diplomové práce. Práce je přiměřeně jazykově i formálně zpracována. Některé obrázky mají špatnou kvalitu.



Univerzita Tomáše Bati ve Zlíně

Fakulta aplikované informatiky

Dotazy k obhajobě

1. Vzhledem k tomu, že se Vám nepodařilo replikovat hodnoty mAP z publikace Chefera (str. 84), jaký to má dopad na celkovou důvěryhodnost kvantitativního srovnání vizualizačních metod ve této práci?
2. Plánujete v budoucnu dále rozvíjet tuto práci?

Celkové hodnocení práce:

Známku uvede oponent dle svého uvážení dle klasifikační stupnice ECTS:

A – výborně, B – velmi dobře, C – dobře, D – uspokojivě, E – dostatečně, F – nedostatečně.

Stupeň F znamená též „nedoporučuji práci k obhajobě“.

Předloženou diplomovou práci doporučuji k obhajobě a navrhuji hodnocení

A - výborně.

V případě hodnocení stupněm „F – nedostatečně“ uveďte do připomínek a slovního vyjádření hlavní nedostatky práce a důvody tohoto hodnocení.

Datum

2.6.2025

Podpis oponenta diplomové práce