

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Analyza kombinování poziční a strukturální informace v grafových neuronových sítích
Jméno autora:	Adéla Schwanzerová
Typ práce:	bakalářská práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra matematiky
Vedoucí práce:	Pavel Procházka
Pracoviště vedoucího práce:	CISCO SYSTEMS, Karlovo náměstí 2097/10, 120 00, Praha

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání a motivace k jeho vypsání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce a krátké průvodní slovo k motivaci pro zadání práce</i>	
<p>Bakalářská práce navazuje na výzkumnou aktivitu v rámci našeho týmu, kde jsme zkoumali výkonnost GNN v závislosti na stupni zjednodušení grafu, tj. redukce jeho velikosti. Výstupem tohoto výzkumu byla konkrétní metoda vedoucí na zjednodušený graf, který vznikl z původního grafu sloučením (klastrováním) jeho uzlů. GNN na tomto zjednodušeném grafu je časově i paměťově méně náročná za cenu snížení přesnosti. Důvodem snížení její přesnosti je nemožnost rozlišit uzly v rámci jednoho klastru, protože GNN predikce je aplikovaná na celý klastr. Motivace pro zadání bakalářské byla eliminovat toto omezení tím, že v rámci klastru budeme uvažovat další model (např. GNN), který umožní rozlišit uzly v rámci jednoho klastru. V zadané práci proto uvažujeme 2 typy modelů. První model zpracovává celý klastr jako jeden uzel v grafu a druhý model je aplikován na grafu pro každý klastr. Existuje celá řada možností kombinování informací z těchto dvou modelů. Základním přístupem pro zadání práce je zvolen jednoduchý součin predikcí modelů s interpretací násobení apriorní a posteriorní pravděpodobnosti.</p> <p>Za nejobtížnější část zadání považuji pochopení principů GNN a naší předchozí práce, na kterou zadání navazuje. Praktická část poté spočívá v provedení experimentů pomocí připravené implementace a následné vyhodnocení výsledků těchto experimentů.</p>	

Splnění zadání	splněno s výhradami
<i>Posudte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
<p>Předložená práce obsahuje všechny body zadání. Rešeršní část považuji za splněnou. Výsledky praktické části jsou prezentovány ve formě několika grafů. Odpovídající interpretace těchto výsledků a jejich vyhodnocení ovšem buď zcela chybí nebo jsou popsány velmi nepřesně.</p>	

Aktivita a samostatnost při zpracování práce	průměrná
<i>Posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven. Posudte schopnost studenta samostatné tvůrčí práce.</i>	
<p>Studentka navazovala na předchozí spolupráci s naší společností. V rámci řešení bakalářské práce probíhaly konzultace, kde jsme iterovali nad průběžnými výsledky. Vzhledem k výzkumné povaze zadání práce byly jednotlivé kroky postupně upřesňovány a konzultovány v závislosti na aktuálních výsledcích. Bohužel část pokynů vyplývajících z těchto konzultací zůstalo nepochopeno, což se ukázalo v některých částech odevzdané práce.</p>	

<p>Odborná úroveň <i>Posudte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i></p>	<p>podprůměrná</p>
<p>Předložená práce trpí řadou faktických nepřesností, které zřejmě pramení z nepochopení dané problematiky v odpovídající hloubce. Přestože práce obsahuje některé relevantní postřehy, jejich prezentace je nedostatečná a z textu těžko pochopitelná. Závěr práce neodpovídá na zadání práce. Tvzení v závěru, že se nepovedlo překonat výsledky z referenční práce (vycházející zřejmě z Obr. 10) považuji za neprůkazné. Za prvé Obr. 10 zobrazuje rozdílný výsledek logistické regrese, což nedává smysl, jelikož je tento model nezávislý na daném grafu. Za druhé by to znamenalo, že GNN aplikovaná na jednotlivé klastry nepřináší žádnou informaci, což neplatí minimálně v extrémním případě, kdy uvažujeme jako klastr celý graf. Dalším problémem je absence vysvětlení a motivace některých parametrů použitých v evaluaci, např. proč jsou porovnány redukce s různým parametrem multistep? Sekce 7.2 zmiňuje použití metody Naive Bayes, ale vůbec není vysvětlený kontext, na co se Naive Bayes používá. Zřejmě byl přejat parametr s názvem naive v konfiguraci experimentu, který ovšem nemá žádnou souvislost s Naive Bayes metodou. V neposlední řadě chybí vysvětlení, jaký je vztah mezi logistickou regresí a dekompozicí grafu.</p>	
<p>Formální a jazyková úroveň</p>	<p>průměrná</p>
<p><i>Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku.</i></p> <p>Přestože v teoretické části je formálně zavedeno mnoho pojmů definic, praktická část je vedena pouze slovně s často vágním vyjadřováním a použitím pojmů, které vysvětleny nebyly. Dále práce obsahuje několik překlepů a drobných typografických problémů. Např. chybí reference v textu na snipety kódu a na Obr. 10, některé citace jsou špatně formátované např. [65], odstavec 7.2.2 končí nedokončenou větou, obrázky a grafy by mohly být přehlednější a lépe popsáné, rozlišení některých obrázků by mohlo být vyšší (případně použit vektorový formát).</p>	
<p>Výběr zdrojů, korektnost citací</p>	<p>výborné</p>
<p><i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posudte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i></p> <p>Rešerše literatury je nejsilnější stránka práce. Práce zasazuje do kontextu 66 relevantních zdrojů pokrývajících zkoumanou oblast.</p>	
<p>Další komentáře a hodnocení</p>	
<p><i>Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.</i></p> <p>Předložená práce měla posloužit jako podklad k našemu dalšímu výzkumu v oblasti GNN a její aplikaci na problémy kybernetické bezpečnosti. Z tohoto pohledu práce přináší velmi málo nových poznatků. Zásadní poznatek, že navržené kombinování poziční a strukturální informace, nevede k očekávanému výsledku, lze vyčíst z uvedených grafů. Ačkoliv samotný fakt, že navržená metoda nevede k očekávaným výsledkům, je ve výzkumu akceptovatelný, zcela chybí analýza, která by přinášela odůvodnění proč a za jakých podmínek se tak děje, přestože návrh této analýzy byl probrán na konzultacích.</p>	

III. CELKOVÉ HODNOCENÍ A NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Případně uveďte otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Pozitivně hodnotím rešeršní část práce, kde studentka prokázala schopnost pochopit a logicky začlenit velkou část relevantní literatury. Bohužel, prezentace zejména praktické části sráží kvalitu odevzdané práce. Negativně také hodnotím fakt, že studentka nevyužila dostupný prostor v rámci konzultací pro pochopení souvislostí.

Při obhajobě se doporučuji zaměřit na srozumitelné vysvětlení cíle práce a jak provedené experimenty na tyto cíle odpovídají.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **D - uspokojivě**.

Datum: 11.1.2023

Podpis:

