POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	**Representation Learning and Adversarial Sample Generation for Strings**
Jméno autora:	**Marek Galovič**
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Doc. Ing. Václav Šmídl, Ph.D.
Pracoviště oponenta práce:	Katedra počítačů, FEL, ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání — **mimořádně náročné**

Hodnocení náročnosti zadání závěrečné práce.

Zadání kombinuje nejnovější metody z oblastí reprezentace pravděpodobnosti a adverzariálního učení, považuji jej za mimořádně náročné.

Splnění zadání — **splněno**

Posuďte, 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.

Zadání bylo splněno, množství odvedené práce je nadstandardní.

Zvolený postup řešení — **vynikající**

Posuďte, zda student zvolil správný postup nebo metody řešení.

Autor se seznámil s metodami používanými pro reprezentaci neurčitosti, kde dokonce navrhl vlastní rozšíření. Dále aplikoval metody optimalizace pro učení s protivníkem na velmi aktuální a náročná reálná data. K postupu řešení nemám co vytknout.

Odborná úroveň — **A - výborně**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

Vzhledem k unikátnosti dat a použití nejnovějších technik je výsledek práce ojedinělý a zajímavý pro praktické použití.

Formální a jazyková úroveň, rozsah práce — **B - velmi dobře**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Práce je psána v angličtině velmi dobré úrovně s minimem překlepů a nejasností. Uvítal bych pouze podrobnější popis experimentů, v některých částech práce není úplně jasné o jaké výsledky se jedná. Souhrnný popis experimentálního protokolu by usnadnil čtení, v současné podobě je třeba spoustu souvislostí domýšlet a dohledávat. Například v tabulkách 4.14 a dále není uvedeno jaké hodnoty zobrazují, je třeba dohledat že success rate. Rozsah práce odpovídá obvyklé délce bakalářské práce, je však zřejmé, že autor provedl více experimentů než je popsáno v práci. Popisky os u grafů by měly být větší, fonty by velikostí měly odpovídat velikosti písma v hlavním textu práce.

Výběr zdrojů, korektnost citací — **A - výborně**

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, 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.

Teoretická část práce je založena na velmi dobrém průzkumu literatury a práce jsou správně citovány. Jedinou výhradu mám ke slovu „novel“ u hodnocení attention mechanismu pro agregaci bagů. Tento přístup je dnes již klasický, viz [1] a následné práce.

[1] Ilse, Maximilian, Jakub Tomczak, and Max Welling. "Attention-based deep multiple instance learning." In *International conference on machine learning*, pp. 2127-2136. PMLR, 2018.

Další komentáře a hodnocení

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.

Práci považuji za velmi zdařilou. Pokud by se vylepšila experimentální část, myslím, že by bylo možné výsledky publikovat.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

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

Práci považuji za velmi dobrou mám jen pár dotazů k experimentům:
1. První zmínka o výsledcích adversarial attack je v části 4.3, kde je citlivostní studie na parameter beta v tabulce 4.1. Není zřejmé, za jakých podmínek byl experiment prováděn. Jak se vztahuje tato tabulka k dalším výsledkům uváděným později?
2. Výsledky pro LBA v tabulce 4.2 jsou sice lepší než pro mean+max, ale jen nepatrně a jen pro dostatečně vysoký počet hlav. Považujete za prokázané, že vhodné LBA používat? Například s ohledem na nevýhody LBA.
3. Ze srovnání metod v části 4.5 vyplývá významná citlivost na zvolené parametry metody. Výsledky jsou prezentovány jen na relativně hrubé mřížce. Pro účely BP to považuji za dostatečné. Můžete stručně zhodnotit svůj dojem, jestli závisí výsledky skutečně na metodě nebo spíše na volbě jejích parametrů?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm A - výborně.

Datum: 6. června 2021 Podpis: