

I. IDENTIFIKAČNÍ ÚDAJE

| | |
|-----------------------------------|--|
| Název práce: | Odhad rychlosti na segmentech silniční sítě |
| Jméno autora: | Martin Koryťák |
| Typ práce: | bakalářská |
| Fakulta/ústav: | Fakulta elektrotechnická (FEL) |
| Katedra/ústav: | Katedra Kybernetiky |
| Oponent práce: | Ing. Mgr. Vladimír Kunc |
| Pracoviště oponenta práce: | Katedra počítačů (FEL) |

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

| | |
|---|-------------------|
| Zadání | náročnější |
| <i>Hodnocení náročnosti zadání závěrečné práce.</i> | |
| Zadání považuji za náročnější, neboť cílem práce je analyzovat modely odhadující rychlost vozidla na segmentech silniční sítě bez dalších limitací na vlastnosti modelů – je třeba prozkoumat relativně širokou množinu tříd rozdílných modelů a následně navrhnout řešení využívající dopravní data ke zpřesnění odhadů. | |

| | |
|--|----------------|
| Splnění zadání | splněno |
| <i>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.</i> | |
| Zadání bylo splněno bez výhrad. Autor prozkoumal hlavní třídy existujících modelů, popsal jejich vlastnosti a aplikace a dále navrhl řešení, které dokázalo využít dopravní data. Toto řešení bylo následně implementováno a experimentálně ověřeno. | |

| | |
|--|-------------------|
| Zvolený postup řešení | vyňikající |
| <i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i> | |
| Postup řešení byl správně zvolen. Řešení obsahuje porovnání několika možných architektur neuronových sítí, výběr vhodného optimalizačního postupu (SGD, RMSprop, Nadam) a výběr vhodných hyperparametrů. Vybrané modely jsou pak srovnány z pohledu hned několika metrik (MAE, MSE, MedAE) a také z pohledu R^2 . Jednotlivé modely jsou také porovnávány na různých datových množinách, z nichž některé obsahují augmentovaná data pro lepší učení. Řešení také obsahuje získání dat a popsání postupu jejich přípravy – tj. stažení obrazových dat pomocí Google API, jejich augmentaci a normalizaci. Autor používá opakované trénování neuronových sítí, což umožňuje lepší odhady chyby – tento postup je preferovaný, třebaže v daném oboru často opomíjený vzhledem k výpočetní náročnosti trénování neuronových sítí (autor propočítal přes 370 dní procesorového času). | |

| | |
|---|--------------------|
| Odborná úroveň | A - výborně |
| <i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i> | |
| Práce je z hlediska odbornosti na vysoké úrovni, obsahuje poznatky z literatury a koncepty, které jdou nad rámec studia. Práce využívá state-of-the-art konvoluční sítě. Práce je jedinečná také svým zaměřením na pražskou dopravní síť. | |

| | |
|--|--------------------|
| Formální a jazyková úroveň, rozsah práce | A - výborně |
| <i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i> | |
| Práce je psána v českém jazyce, je psána jasně a přehledně. Například názvy proměnných z kódu a souborů jsou v textu odlišeny písmem s pevnou šířkou, matematické proměnné a zvláštní názvy kurzivou, díky čemuž se práce příjemně čte. Drobnou výtku mám k použití kurzivy v textu u jednotek, což by mělo být vyhrazeno pouze pro matematické proměnné. Formální zápis je také vždy správný. Rozsah práce je na průměrné úrovni, velkou část tvoří diagramy. Některé sekce by mohly být pojmenovány lépe – např. sekce 2.4.3.4 Dropout by měla být spíše pojmenována Regularizace, protože se nezabývá jen regularizační metodou dropout, nýbrž popisuje i L1 a L2 regularizaci. U grafů v přílohách by bylo lepší, kdyby nezobrazovaly extrémní hodnoty na počátku učení. | |

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.

Výběr zdrojů je adekvátní bakalářské práci. Nicméně zdrojů mohlo být využito více, v některých částech se zdá, že student citoval zdroje spíše namátkou (např. v sekci 2.5 Hluboké učení z leteckých a satelitních snímků jsou popsány jen některé články, třebaže v dané oblasti existuje velké množství článků – bylo by lepší na některé odkázat). Autor cituje korektně a vlastní práce je vždy odlišena od převzatých prvků, k porušení citační etiky nedochází. Třebaže autor jmenuje použité knihovny, bylo by lepší, kdyby citoval whitepapers, pokud daná knihovna ho má (např. Tensorflow či Scikit-learn).

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.

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.

Práce je kvalitně zpracována, velmi pozitivně hodnotím použití různých diagramů, které usnadňují pochopení prezentovaných konceptů. Co se týče technické stránky zpracování práce, nemám výhrad. Zejména oceňuji použití vektorové grafiky pro jednotlivé diagramy a grafy, neboť to umožňuje libovolné přiblížení při prohlížení ve formátu PDF.

Drobné výtky mám k popisu metodologie – práce využívá předučené sítě ResNet a DenseNet na datové množině ImageNet, nicméně tyto sítě se používají ke klasifikaci, chybí mi proto popis způsobu, jakým byla síť upravena pro regresní úlohy (tj. jak vypadají horní vrstvy), třebaže nějaký náznak je v tabulce 4.2. Dále chybí informace, která chybová metrika byla použita k optimalizaci sítě (zda MAE nebo RMSE), práce pouze říká, že nejčastěji se používají obě metriky, ale že MAE je preferovaná pro menší citlivost na extrémní hodnoty.

Dále oceňuji optimalizaci hyperparametrů. Nicméně by bylo lepší, kdyby byla tato optimalizace lépe popsána – není zcela jasné, jak jsou hyperparametry optimalizovány (např. z jakého intervalu byly vybírány) a kolikrát byla síť vyhodnocována v různých parametrizacích.

Práce porovnává jednotlivé datové množiny a různé způsoby trénování (finetuning, apod.), nicméně neobsahuje tabulku ani grafy, které by srovnávaly výsledky (ani z textu se není možné dočíst, jak se lišily výsledné chyby), práce pouze obsahuje grafy v příloze, kde jsou jednotlivé chyby srovnány podle architektur – srovnání datových množin a způsobů trénování téměř nelze z těchto grafů vyčíst. Naopak chválím, že práce obsahuje tabulku srovnávající různé metriky několika nejlepších sítí místo popisu pouze jedné nejlepší sítě.

V práci se píše „Po naučení těchto vrstev se odmrazila část vah state-of-the-art sítě a proběhlo jejich doučení“, zajímalo by mě, jak se vyvíjela chybová metrika během učení zmrazených vah a po odmražení. Došlo po odmražení vah sítě k rapidnímu zhoršení přesnosti? Kolik vrstev bylo odmražováno?

Celkově práce působí kvalitním dojmem jak po technické, tak po odborné stránce. Práce se příjemně čte, všechny použité koncepty jsou jasně a přehledně vysvětleny.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **A - výborně**.

Datum: 30.5.2018

Podpis: