

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Predikce rychlostí úseků cyklistických tras pomocí strojového učení
Jméno autora:	Zemek Alexandr
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Doc. Ing. Michal Jakob, Ph.D.
Pracoviště oponenta práce:	Katedra počítačů, FEL ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání <i>Hodnocení náročnosti zadání závěrečné práce.</i>	Náročnější
Zadání považuji za náročnější - student se musel v práci netriviálním způsobem zpracovat data různých formátů (zejména reálná GPS záznamy cyklistických trajektorií a mapové podklady ve formátu OpenStreetMap) a na zpracovaná data následně aplikovat prediktivní modely strojového učení.	

Splnění zadání <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>	Splněno s výhradami
Při návrhu a implementaci prediktivního modelu se dle zadání měl student soustředit na metody strojového učení "multiple-instance learning". Tento druh metod je však v práci popsán pouze v závěrečné kapitole při popisu možných budoucích rozšíření práce. Implementované metody popsané v práci pracují s pouze fixní délkou vektoru příznaků. Z textu není zřejmé proč nedošlo k implementaci metod "multiple-instance learning" a k experimentování s variabilní délkou vektoru příznaků.	

Zvolený postup řešení <i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	B - velmi dobře
Zvolený přístup k řešení daného problému považuji za adekvátní. Student využil softwarové nástroje a knihovny, které se ve výzkumné praxi k řešení podobných problému využívají. Menší výhrady mám jediné k zvolení učení modelu doby jízdy na segmentech silniční sítě místo rychlosti. Výsledné hodnoty modelu jsou přímo závislé na délce segmentu, což potvrdilo porovnání významnosti příznaků. Význam příznaku délky pak zastiňuje ostatní příznaky, které jsou pro odhad rychlosti cyklisty důležitější, zejména převýšení nebo povrch.	

Odborná úroveň <i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	B - velmi dobře
Odborná úroveň práce je na dobré úrovni. Student v práci detailně metody neuronových sítí a jejich využití pro řešení zadaného problému. V práci mi nicméně chybí samostatná kapitola, která by detailně formalizovala řešený problém a jeho součástí. Kladně naopak hodnotím aplikaci softwarově-inženýrských postupů, kdy student zdárně implementoval projekci surových GPS sekvencí na podkladový graf silniční sítě pro vytvoření trénovacího a testovacího datasetu, včetně obohacení o dodateční data z komplexních mapových podkladů OpenStreetMap.	

Formální a jazyková úroveň, rozsah práce <i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	B - velmi dobře
--	------------------------

Text je psaný anglicky na slušné úrovni. Občas se v textu vyskytují překlepy, které ale nenarušují srozumitelnost textu. V práci je použito vhodné množství obrázků a tabulek, které doplňují text. Celková čitelnost a přehlednost textu bohužel není dobrá a trpí jednak ne zcela vhodným členěním textu do kapitol a chybějící detailní formalizací problému, na kterou by se zbytek práce odkazoval. Ne zcela vhodné členění strukturování textu do kapitol ilustruje čtvrtá kapitola "Solution Approach", kde je současně popisovaná formální definice řešení problému, implementace a evaluace modelů. Celkový rozsah textu je standardní.

Výběr zdrojů, korektnost citací

B - velmi dobře

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.

Bibliografické údaje jsou úplné a v souladu s citačními zvyklostmi. Avšak v části textu, která se věnuje relevantní literatuře bych očekával detailnější rozbor většího počtu článků věnujících se predikci času jízdy se zaznamenaných sekvencí GPS pozic. Sám student v první větě této kapitoly naznačuje, že studií věnujících se této problematice existuje několik ("There have been many studies about predicting velocities not just for bicycles, but for the traffic flow in general."), aniž by je v rešerši analyzoval.

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.

V práci student popisuje filtraci segmentů s naměřenou rychlostí cyklistů vyšší než 45 km/h. Rád bych se dozvěděl, proč si student určil tuto hraniční rychlost a jestli neuvažoval o jiné filtraci, jako například na základě průměrné rychlosti celé trasy. Segmenty silniční sítě, kde cyklista může bez větších problémů dosáhnout rychlost vyšší než definovaná hranice existují. Jedná se například o segmenty s klesáním při sjezdu s kopce.

Příznak "Traffic density", který byl získaný z OSM dat a který měl poměrně vysokou důležitost není v textu detailněji popsán. Jak přesně byl tento příznak získán? Jaké typy silnic byly použity při jeho výpočtu?

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.

Student v práci nastudoval metody strojového učení pomocí neuronových sítí, implementoval několik variant prediktivních modelů, které následně otestoval na reálných datech, které netriviálním způsobem předzpracoval. Studentovi se sice nepodařilo aplikovat metody multiple instance learning specifikované v zadání, ale i tak je zřejmé, že odvedl na řešení velký kus práce. Celkové vyznění práce sráží text, kterému chybí přesnější formalizace řešeného problému a celkově kvalitnějším zpracování.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 8. 6. 2020

Podpis: