

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Využití regresní analýzy sportovního výkonu pro automatizované dolování dat
Jméno autora:	Bc. Vojtěch Zahradník
Typ práce:	diplomová práce
Fakulta:	Fakulta jaderná a fyzikálně inženýrská (FJFI)
Katedra:	Katedra KSI
Oponent práce:	Ing. Jan Jirka
Pracoviště oponenta práce:	ADASTRA, s.r.o.

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>V rámci diplomové práce si vážím náročnosti zpracování dat, zejména při práci s nestandardními formáty FIT a GPX. Tato část práce vyžadovala od autora nejen technickou zručnost, ale také schopnost interpretovat data z těchto formátů, vyžadující znalost i v oblasti datové problematiky. Autor pro dosažení svých cílů využil heuristických algoritmů pro hledání optimálních parametrů v modelu, což představuje náročnou aplikaci těchto metod, která není triviální. V rámci zadání bylo navíc požadováno vytvoření uživatelského reportu pro prezentaci výsledků modelu, což znamenalo propojení technického aspektu práce s uživatelsky srozumitelným výstupem. Tato kombinace technických dovedností a pochopení uživatelské perspektivy zdůrazňuje komplexnost a náročnost celého projektu.</p>	

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>	
<p>V rámci práce bylo stanoveno pět klíčových cílů. Prvním z nich byla optimalizace modelů regresní analýzy, což bylo úspěšně dosaženo prostřednictvím optimalizace hyperparametrů pomocí heuristik. Druhým cílem bylo dosažení přesných předpovědí celkových časů závodu, což bylo úspěšně implementováno ve výsledných modelech na základě sportovních aktivit. Třetím cílem byla rešerše podobných aplikací, kde autor kriticky zhodnotil pozitivní a negativní aspekty daných aplikací. Čtvrtým cílem byla automatizace zpracování dat pro generování interpretovatelného HTML reportu s výsledky modelu, poskytujícího relevantní informace pro běžce. Pátým a posledním cílem bylo vyhodnocení finální aplikace na reálných závodech, kde autor aplikoval své výzkumné poznatky v praxi během vlastních běžeckých závodů a představil výsledky svého modelu.</p>	

Zvolený postup řešení	vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
<p>Po pečlivém prostudování předložené diplomové práce a provedení analýzy zvoleného postupu řešení docházím k závěru, že student přistoupil k řešení svého tématu adekvátně. Zvolené metody byly vhodně vybrány s ohledem na specifika tématu a stanovené cíle práce. Student se prokázal schopným zdůvodnit výběr konkrétních postupů a metod, což přispívá k celkové kvalitě provedené práce.</p>	

Odborná úroveň	průměrná
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
<p>Práce je celkově dobře strukturovaná, kde většina textu navazuje na sebe logicky a plynule. Výjimkou je kapitola "Heuristik", která by měla být zařazena namísto kapitoly o výzkumu trhu. Tím by se zajistila lepší návaznost teorie strojového učení, běhání a následné implementace. Dále by bylo vhodné umístit kapitolu</p>	

"Feature Engineering" (5.5) před kapitolu o ladění hyperparametrů (5.4), neboť v této části autor odvozuje nové proměnné, které jsou následně brány v úvahu v kapitole o ladění hyperparametrů.

Významnější výhrady lze vznést vůči grafickým vizualizacím v rámci práce. Obrázky by měly obsahovat informace o jejich původu – zda byly vytvořeny autorem nebo převzaty z jiného zdroje. Kromě toho je třeba upozornit na absenci jednotek u os v grafikách v kapitolách 5 a 6, což omezuje srozumitelnost prezentovaných výsledků.

Celkově by drobné úpravy v uvedených oblastech mohly významně přispět ke zlepšení kvality prezentace a srozumitelnosti obsahu diplomové práce.

Formální a jazyková úroveň

průměrná

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

Chválím autora za zpracování práce v anglickém jazyce, jehož kvalita dosahuje vysoké úrovně. Přestože v rámci některých odstavců dochází k opakujícím se spojkám, tato odchylka zřídka upoutá pozornost čtenáře, s výjimkou situací, kdy se aktivně věnuje identifikaci možných chyb a nedostatků v obsahu práce.

Výběr zdrojů, korektnost citací

výborné

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.

Autor efektivně využívá kvalitních zdrojů, které jsou klíčové pro zpracování diplomové práce. Většinou dodržuje správná pravidla citování; avšak občasné nejasnosti vztahující se k obsahu citací se vyskytují. Konkrétně citace číslo 59 v oddílu 1.1 pravděpodobně odkazuje i na podkapitoly uvedené v rámci tohoto oddílu, které jsou však bez citací. V budoucnu by autor měl zajistit, aby citace byly explicitně spojeny i s odpovídajícími podkapitolami.

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.

Práce plně odpovídá mému očekávání. Pozitivně hodnotím zvolené metody a technické provedení práce. Nicméně, některé nedostatky se objevily v rámci logické návaznosti textu a formálních chyb. Zlepšení by bylo dosaženo rozšířením práce o další informace týkající se doménových znalostí běhání a možného přesahu do oblasti podnikání.

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 psána kvalitní angličtinou a obecně dobře strukturovaná, avšak existují některé formální body, které vyžadují úpravy. Nejvíce oceňuji originální téma, náročné zpracování a získání dat, zejména při práci s nestandardními formáty FIT a GPX. Tato část vyžadovala od autora technickou zručnost a schopnost interpretovat a získat data pro vytvoření práce.

Vzhledem k identifikovaným výhradám a otázkám pro obhajobu je nutné, aby autor zodpověděl následující otázky před komisí:

Otázky:

- 1) Která data by mohla obohatit stávající model a co nejvíce ovlivňuje běžce před závodem?
- 2) Proč jste vytvářel predikce den předem, a ne přímo před závodem?
- 3) Jak byste využil výsledky modelů během závodu s ohledem na současnou technologii nabízet dynamický alerting při nedodržování predikovaných metrik?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

Datum: 4.2.2024

Podpis: 