



Hodnocení vedoucího závěrečné práce

Student: Filip Kojan
Vedoucí práce: Ing. Karel Klouda, Ph.D.
Název práce: Predikce výsledků zápasů v NHL
Obor: Znalostní inženýrství

Datum vytvoření: 10. 6. 2019

Hodnotící kritérium:	Způsob hodnocení – následující škálou 1 až 4:
1. Splnění zadání	1=zadání splněno, 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
Popis kritéria: Posuďte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posuďte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.	
Komentář: Se zadáním přišel autor sám. Predikce výsledků sportovních utkání je celkem prozkoumané téma, novinkou mělo být využití tzv. moderních hokejových statistik, které nestojí jen na "řídých" jevech jako jsou góly a výsledky zápasů. Predikovat výsledky zápasů v NHL je složitější než ve většině vrcholných soutěžích kolektivních sportů, protože NHL má v sobě zabudované mechanismy (platové stropy, přidělování míst v draftu), které zabraňují tomu, aby se od sebe týmy výkonnostně příliš vzdalovaly. Proto bylo jasné, že úspěšnost predikce získaných modelů se bude držet mezi 54% úspěšností základního modelu "vyhraje domácí tým" a úspěšností sázkových kanceláří, která lehce přesahuje 60 %. Autor se se svými výsledky dokázal těmto 60 % přiblížit, alespoň na testovacích datech. Na úvod posudku bych ještě rád zmínil jednu okolnost, která ovlivňuje mé hodnocení. Pan Kojan měl totiž vážný úraz zhruba měsíc před odevzdáním práce a odevzdaná práce tedy odpovídá zhruba tomu, co bylo hotové v tuto dobu. Nemohl tak dotáhnout většinu dalších nápadů, které jsme měli, a které by z práce udělaly lepší dílo. Takto jde pouze o práci základním způsobem splňující zadání. Vzhledem k tomu, že pan Kojan na práci pracoval velice usilovně a již dlouho dopředu měl tyto základní výsledky a i část textu, věřím, že nebýt úrazu, dotáhl by práci výrazně dál. Snažím se sice hodnotit skutečný stav odevzdané práce, přesto se asi neubráním tomu, abych trochu přimhouřil oko, neboť vím, že pan Kojan patřil k nejpilnějším studentům, kterým jsem kdy vedl závěrečnou práci.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	75 (C)
Popis kritéria: Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti. Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.	
Komentář: Práce je psána srozumitelně a myslím, že celkem přehledně popisuje výsledný model a dataset. Závěrečné kapitoly (psané s jednou rukou v sádře) už jsou slabší a je na nich vidět spěch a snaha práci nějak dokončit. Popis experimentů tak není úplně přehledný a pro čtenáře neznajícího kontext asi místy nesrozumitelný.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	80 (B)
Popis kritéria: Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výkumná a experimentální práce – opakovatelnost experimentů	

Komentář:

Jak to tak bývá, nejvíce práce dala příprava dat. NHL poskytuje veřejné (a nezdokumentované) API, přes které jsou přístupná velmi podrobná data o jednotlivých zápasech i týmech. Autorovi se podařilo tato data zpracovat a spočítat z nich pomocí skriptu v jazyce Python pokročilé statistiky jako jsou CORSI a jiné. Díky tomu mohl tyto statistiky použít v machine learningových modelech využitých pro predikci výsledků zápasů. Ne všechna data byla ovšem dostupná přes toto API, bylo tak třeba dát dohromady více zdrojů a vznikl tak dataset, který patří k nekompletnějším datasetům dostupným veřejnosti. Autor také provedl standardní procedury vedoucí k vyladěnému modelu, který predikuje výsledky zápasů s relativně vysokou úspěšností.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

4. Hodnocení výsledků, jejich využitelnost

80 (B)

Popis kritéria:

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Komentář:

Výsledný model nedosahuje úspěšnosti, která by nějak vynikala např. nad modely sázkových kanceláří (což by bylo až senzační). Užitečným výstupem je výsledný dataset a skripty, které jej dovolují kontinálně doplňovat o nové zápasy. Neméně užitečným mohou být pro zájemce o práci s daty z NHL popisy datových zdrojů a NHL API.

Hodnotící kritérium:

Způsob hodnocení – následující škálou 1 až 5:

5. Aktivita a samostatnost studenta

5a:

1=výborná aktivita,
2=velmi dobrá aktivita,
3=průměrná aktivita,
4=slabší, ale ještě dostatečná aktivita,
5=nedostatečná aktivita

5b:

1=výborná samostatnost,
2=velmi dobrá samostatnost,
3=průměrná samostatnost,
4=slabší, ale ještě dostatečná samostatnost,
5=nedostatečná samostatnost

Popis kritéria:

V souvislosti s průběhem a výsledkem práce posuďte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven (5a). Posuďte schopnost studenta samostatně tvůrčí práce (5b).

Komentář:

Mám-li hodnotit aktivitu a samostatnost autora před úrazem (viz výše), musím obojí hodnotit jako výborné. Kritizovat autora za nechození na konzultace, když vím, že leží na jednotce intenzivní péče se mi přičí.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

6. Celkové hodnocení

82 (B)

Popis kritéria:

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.

Text hodnocení:

Výsledné hodnocení se mi určuje těžko. Bez kontextu bych se asi klonil ke známce C, nakonec jsem se ale rozhodl vzhledem ke všemu výše uvedenému hodnotit práci známkou B.

Podpis vedoucího práce: