



Hodnocení vedoucího závěrečné práce

Vedoucí práce: Ing. Jaroslav Kuchař, Ph.D.
Student: Bc. Martin Macho
Název práce: Využití strojového učení a AI pro extrakci informací z webu
Obor / specializace: Znalostní inženýrství
Vytvořeno dne: January 18, 2022

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Cílem ZP je seznámit se problematikou extrakce informací z webu a aplikovat v doméně extrakce informací z produktových stránek e-shopů. Zvolený přístup je postaven na detekci objektů v obrázkové reprezentaci dané stránky. Jedná se o kombinaci dvou oblastí - webového a znalostního inženýrství.

2. Písemná část práce

89 /100 (B)

V písemné části jsem nezjistil žádné zásadní nedostatky. Práce obsahuje všechny důležité a informačně bohaté část. Po věcné, jazykové i obsahové stránce také neshledávám žádný problém. Po typografické stránce by bylo např. vhodné sladit umístění obrázků k textu apod. Text a obsah je v souladu s citačními zvyklostmi a normami, stejně tak s licenčními podmínkami.

3. Nepísemná část, přílohy

89 /100 (B)

V rámci práce vzniklo několik dílčích nástrojů pro různé fáze - od extrakce obsahu a obrázků stránek po použití modelů postavených na neuronových sítích. Výsledek je plně použitelný a to bylo ověřeno experimenty. Výtkou z mé strany je horší dokumentace celku a jednotlivých nástrojů. Zejména z pohledu navázání na vzniklé řešení.

4. Hodnocení výsledků, jejich využitelnost

95 /100 (A)

Výstupem práce je množina nástrojů a postupů jak aplikovat detekci objektů v obrázku pro extrakci informací z webu. Velmi důležitou součástí je tvorba datasetu, která

vyžadovala velké úsilí. Zejména implementace a postupy spojené s řešením problémů na velkém množství různých e-shop řešení (375 e-shopů). Navržená metoda využívá aktuální metody pro detekci objektů v obrázcích a rozšiřuje o extrakci z HTML. Získané výsledky odpovídají komplexnosti problému. Celkově má práce potenciál dále se rozvíjet i do výzkumných oblastí.

5. Aktivita studenta

- ▶ [1] **výborná aktivita**
- [2] velmi dobrá aktivita
- [3] průměrná aktivita
- [4] slabší, ale ještě dostatečná aktivita
- [5] nedostatečná aktivita

výborná aktivita

6. Samostatnost studenta

- ▶ [1] **výborná samostatnost**
- [2] velmi dobrá samostatnost
- [3] průměrná samostatnost
- [4] slabší, ale ještě dostatečná samostatnost
- [5] nedostatečná samostatnost

výborná samostatnost

Celkové hodnocení

90 /100 (A)

Student v rámci ZP zpracoval zadané téma. Podrobně se seznámil s problematikou. Navrhl, implementoval a vyhodnotil vhodné řešení. Kvalita výstupů je na vysoké úrovni a je využitelná pro další práce v této oblasti. Vzhledem k uvedenému hodnotím A a doporučuji k obhajobě.

Instrukce

Splnění zadání

Posudte, 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, posudte 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.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, 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. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 52/2021, článek 3.

Posudte, 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.

Nepísemná část, přílohy

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, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

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.

Aktivita studenta

V souvislosti s průběhem a výsledkem práce posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven.

Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posudte schopnost studenta samostatně tvůrčí práce.

Celkové hodnocení

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.