

Posudek oponenta bakalářské práce

Student: Kanda Rastislav
Téma: Automatizovaná extrakce informací z emailů (id 22028)
Oponent: Zbořil František, doc. Ing., Ph.D., UITS FIT VUT

- 1. Náročnost zadání** průměrně obtížné zadání
Zadání považuji za průměrně obtížné. Mělo dojít jak k prostudování dostupných metod a technologií, které by umožnily vytvořit systém pro extrakci informací z emailů, tak k jeho implementaci a ověření na reálných datech.
- 2. Splnění požadavků zadání** zadání splněno
Zadání bylo splněno a funkčnost výsledného systému byla ověřena.
- 3. Rozsah technické zprávy** je v obvyklém rozmezí
- 4. Prezentační úroveň předložené práce** 80 b. (D)
Práci vytýkám hlavně to, jak vysvětluje problém a jeho řešení. Připadá mi, že je zbytečné věnovat tolik prostoru známým věcem, jako jsou HTTP, HTML atd. a aby se na vlastní přínos studenta dostalo až ke konci práce, tj. od strany 27. Navíc text není příliš jasný. Například pro čtenáře, který nezná algoritmus ROBULA+, musí být podivné, když se o algoritmu jako takovém v kapitole, která jej má představovat (kap. 3) mnoho nedozví. Algoritmus je uveden až na straně 31. Za menší prohrěšek považuji to, že některé informace se v práci zbytečně opakují, například že NLP je Natural Language Processing apod.
- 5. Formální úprava technické zprávy** 75 b. (C)
Text je psán slovensky a jazykově se mi zdá v pořádku. Ani po typografické stránce nemám k textu vážnější připomínky a celkově se jedná o průměrný text co se formální stránky týče.
- 6. Práce s literaturou** 80 b. (B)
Bez jednoho třiceti zdrojů, jak už bývá obvyklé u těchto prací vesměs z internetu, je v práci řádně citováno a pokrývá tématicky vhodně předložené řešení.
- 7. Realizační výstup** 85 b. (B)
Realizační výstup je funkční. Je schopen správně vyhledat požadované informace ve třech čtvrtinách emailů, což student považuje za zlepšení v porovnání se systémy, které se mu podařilo dohledat. Proto hodnotím tento bod nadprůměrně.
- 8. Využitelnost výsledků**
Projekt měl vyústit v řešení, které by bylo využitelné společností Kiwi. Poměrně uspokojivá schopnost systému vyhledávat požadovaná data v emailech patrně povede k tomu, že výsledná aplikace bude využívána pro komerční účely.
- 9. Otázky k obhajobě**
 - Na straně 20 píšete, že spolehlivost 51.9% při zpracování emailů je přijatelná. Na jakém základě jste toto hodnocení udělal a kde je hranice přijatelnosti a jak s určí?
 - Můžete konkrétněji vysvětlit vaše konstatování v sekci 4.7, že "Byť webovou aplikáciou, znamená dokázat zvládnúť niekoľko požiadavkov v jeden moment."?
- 10. Souhrnné hodnocení** 75 b. dobře (C)
Práce má své kladné stránky, splnila zadání i očekávání, které do ní zadavatelé vkládaly. Některé slabší stránky zejména co se týče odevzdaného textu mě vedou k průměrnému hodnocení stupněm C.

Prohlášení: Uděluji VUT v Brně souhlas ke zveřejnění tohoto posudku v listinné i elektronické formě.

V Brně dne: 4. června 2019

Zbořil František, doc. Ing., Ph.D.
oponent