

Strukturovaný posudek oponenta bakalářské práce

Autor/autorka práce: Václav Suda

Název práce: Analýza propojení firem využitím PageRanku

Obsah práce:

Vynikající logická struktura, nadprůměrný obsah i rozsah;
Velmi dobrá logická struktura, odpovídající obsah i rozsah;
Vyhovující logická struktura, obsah i rozsah;
Nevyhovující

Komentář:

Kvalita řešení a dosažených výsledků:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Komentář: Cennější než nakonec dosažené výsledky je použitá metodologie a spojení několika typů dolování z dat v jedné práci (web mining + síťová analýza). Za významné považuji informace o firmách a osobách získané dolováním z webových stránek.

Formální úroveň:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Komentář: Škoda že kapitola Úvod je v podstatě jen stručný obsah práce a Závěr pouze shrnutím prováděných činností. Hodilo by se nějaké uvedení do dané problematiky a její hodnocení z většího nadhledu. Dále uvádím některé překlepy v textu, z nichž nejčastější je „mír centrality“ místo „měr centrality“ na str. 28, 29, 39, 40, 41, 42, 47 a potom „v České Republice“ (má být malé r), „betweenes“ místo „betweenness“, „využitím kterého“ místo lepšího „jehož využitím“, „nastavíme, tak aby“ (čárka má být před aby) na str. 1, „klesání hodnocení“ jako nevhodný překlad „rank sink“ na str. 3, „viz vzorec (4)“ na str. 5 (má být „viz vzorec (5)“), chybějící vzorec v kapitole 2.3.3, „soba“ na str. 10 místo „osoba“, takřka identické odstavce na str. 12 a 15, „v řeší metody“ místo „v řešení metody“ na str. 29, „spousta vrcholů“ místo „mnoho vrcholů“ na str. 40, „nejdeme alespoň“ místo „najdeme alespoň“ na str. 41, „tou jsou“ místo „to jsou“ na str. 41, „rejstirk-firem.kurzy.cz“ místo „rejstrik-firem.kurzy.cz“ na str. 42, „přepne se“ místo „přepneme se“ na str. 47.

Práce s literaturou:

Vynikající; Velmi dobrá; Vyhovující; Nevyhovující

Komentář:

Splnění zadání:

Splněno bez výhrad; Splněno s menšími výhradami; Splněno s většími výhradami; Nesplněno

Komentář: Výsledky prezentované v kapitole 9 bylo jistě těžké získat, ale nelze z nich učinit prakticky žádné závěry, protože jsou nevěrohodné a, jak bakalant sám konstatuje, v podstatě protichůdné. Rozdíly v PageRanku na obr. 15 a dalších jsou mikroskopické a při zřejmě neexistujícím konvergenčním kritériu (alespoň se o něm nikde nic nepíše) nejsou patrně prezentované žebříčky stabilní. Konstatování, že žebříčky pro closeness centrality „nejsou úplné“ (na str. 38) je silný eufemismus, protože ve skutečnosti je to jen zlomek všech výsledků.

Doplňující informace k práci: Bakalant se věnoval složitému tématu důkladně a pečlivě, ale bohužel jsem na přiloženém DVD (mimořádně sotva čitelně označeném) nenalezl zdrojové kódy jeho programů, jen JAR a soubory s příponou CLASS, takže není možné hodnotit jeho programátorskou činnost.

Dotazy k práci:

Vysvětlete prosím, jak jste došel k procentům uvedeným na straně 40 dole.

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 7. 8. 2013

doc. Ing. Dalibor Fiala, Ph.D.

