



Hodnocení vedoucího závěrečné práce

Student: Bc. Juraj Polačok
Vedoucí práce: Ing. Marek Sušický
Název práce: Využití grafových databází pro pokročilou analýzu dat
Obor: Znalostní inženýrství

Datum vytvoření: 2. 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: Zadání bylo splněno. Práce obsahuje rešerši stávajících grafových databází, jejich srovnání a následně vlastní benchmark a doporučení pro vhodnost pro komerční projekt.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
2. Písemná část práce	85 (B)
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: Písemná práce splňuje požadované náležitosti, je logicky strukturovaná a jednotlivé kapitoly mají očekávaný rozsah, až na kapitolu realizace, kde by bylo vhodné detailněji popsat postup použitý při benchmarku. Práce obsahuje několik přklepů, které ale nemají vliv na pochopení obsahu. Zbytečné jsou obrázky 3.3 a 3.4, jelikož nenesou žádnou informační hodnotu. Typograficky práce splňuje požadavky a autor cituje korektně a uvádí zdroje, které jsou relevantní. V rešerši jsou zkoumány licence jednotlivých grafových databází a vybraná je užita v souladu s licenčními podmínkami.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
3. Nepísemná část, přílohy	90 (A)
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, Výzkumná a experimentální práce – opakovatelnost experimentů	
Komentář: Provedená měření jsou opakovatelná a vytvořený kód je čitelný. Vybraná technologie je v souladu s požadavky a má vhodné licenční podmínky.	
Hodnotící kritérium:	Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):
4. Hodnocení výsledků, jejich využitelnost	90 (A)
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:

Výsledek bude po drobných úpravách využitelný v praxi. Grafové databáze skrývají velký potenciál, který zatím zůstává nevyužitý, i díky nezralosti technologií. To je vykoupeno vhodným licencováním.

V tomto roce bychom chtěli na základně výsledku této práce realizovat projekt s ostrými daty u zákazníka z bankovního sektoru.

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:

Během řešení jsme se pravidelně se studentem týdně setkávali a domlouvali se na dalším postupu. Student samostatně řešil komplikace vzniklé při zkoumání grafových databází i benchmarku. Byly dodržovány domluvené termíny i řešeny úkoly.

Hodnotící kritérium:

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

6. Celkové hodnocení

85 (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:

Cíle práce byly splněny. V úvodní rešerši grafových databází jsou uvedeny všechny relevantní databáze včetně detailnějších informací a benchmarků, pokud jsou dostupné. Na základě této rešerše je pak v kapitole 3.4 uděláno shrnutí a vybrána nejvhodnější databáze, JanusGraph. Student zprovoznil technologii na výpočetním clusteru, importoval data a vytvořil generátor dat syntetických. Student se vypořádal s problematickým výpočetním clusterem, který neměl dostatečný výkon na upočítání grafu a zmenšil odpovídajícím způsobem datasety. Následně byly provedeny benchmarky pro PageRank a Clustering, ale zdokumentování benchmarků by mohlo být provedeno detailněji. Např. přidat grafy s využitím paměti a CPU, opakovat měření několikrát, aby se vyfiltroval vliv jiných aplikací běžících na clusteru.

Podpis vedoucího práce: