

Posudek oponenta diplomové práce

Student: Veselovský Martin, Bc.

Téma: Získávání znalostí pro modelování následných akcí (id 20133)

Oponent: Burget Radek, Ing., Ph.D., UIFS FIT VUT

- 1. Náročnost zadání** **obtížnější zadání**
Zadání považuji za mírně náročnější z hlediska studijní etapy. Student se musel detailně seznámit s problematikou datových skladů a získávání znalostí z dat a to nejen teoretické, ale i praktické rovině v kontextu konkrétních datových zdrojů.
- 2. Splnění požadavků zadání** **zadání splněno**
Zadání bylo splněno ve všech bodech.
- 3. Rozsah technické zprávy** **je v obvyklém rozmezí**
Technická zpráva se svým rozsahem blíží spíše dolní hranici obvyklého rozmezí.
- 4. Prezentací úroveň předložené práce** **85 b. (B)**
Technická zpráva pokrývá jak studijní etapu zahrnující problematiku datových skladů a získávání znalostí z dat, tak i praktickou část, která se týká návrhu konkrétního datového skladu a implementace souvisejících nástrojů včetně nástrojů pro klasifikaci. Teoretická část je velmi logicky a přehledně strukturovaná a výborně srozumitelná. Praktická část je místy stručnější a náročnější z hlediska porozumění celkovému konceptu řešení. Pomohlo by více obrázků ilustrujících zejména architekturu celého řešení, schéma databáze implementující datový sklad a podobně.
- 5. Formální úprava technické zprávy** **90 b. (A)**
Po formální stránce je technická zpráva velmi pečlivě zpracována, k jazykové ani typografické stránce nemám připomínky.
- 6. Práce s literaturou** **85 b. (B)**
Seznam použité literatury je přiměřeně obsáhlý, všechny zdroje jsou relevantní a pokrývají celou řešenou problematiku. V textu práce jsou zdroje řádně uváděny.
- 7. Realizační výstup** **96 b. (A)**
Realizačním výstupem je datový sklad vybudovaný nad relační databází PostgreSQL a sada nástrojů implementovaných v jazyce Python, které umožňují jednak plnění datového skladu a zejména pak spouštění klasifikačních úloh nad vybranými vzorky dat. Celé řešení je cíleno na praktické využití v oblasti internetové reklamy a predikce její návratnosti. Za zmínku stojí i do jisté míry originální využití rozhodovacích stromů ke generování doporučení ohledně parametrů zadávané reklamy. Ačkoliv výsledná aplikace nedisponuje grafickým uživatelským rozhraním a je tedy použitelná spíše expertem na danou problematiku, než běžným uživatelem, je plně funkční a produkuje výsledky uvedené v technické zprávě. Kladně hodnotím rovněž provedené testování na reálných, poměrně rozsáhlých datech a experimentální srovnání dalších klasifikačních metod.
- 8. Využitelnost výsledků**
Navržený datový sklad a související nástroje jsou použitelné pro analýzu dat v dané oblasti a podporu rozhodování.
- 9. Otázky k obhajobě**
 - Mohl byste nejlépe graficky znázornit schéma databáze použité pro datový sklad (sněhová vločka) naznačené v kapitole 5.1.4?
- 10. Souhrnné hodnocení** **90 b. výborně (A)**
Pan Veselovský detailně nastudoval problematiku datových skladů, OLAP a získávání znalostí z dat a navrhl komplexní a plně funkční aplikaci pro predikci a analýzu výsledků. Jedná se o poměrně experimentálně orientované téma a realizační výstup považuji z inženýrského pohledu za velmi pěkně navržený. Dojem mírně kazí poněkud stručnější technická zpráva, přesto však vzhledem ke kvalitám realizačního výstupu navrhuji hodnotit práci jako výbornou.

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

V Brně dne: 8. června 2017

.....
podpis