

# Posudek oponenta diplomové práce

**Název práce:** Automatizovaná analýza medicínských dat

**Autor:** Pavel Cihlář

## **Obsah práce:**

Cílem práce bylo seznámení se s meta-datovým modelem RDF, dotazovacím jazykem SPARQL a dalšími používanými technologiemi v projektu MRE. Dále navržení vhodné metody pro analýzu dat z RDF, tvorba popisné charakteristiky a metod pro měření závislostí a testování hypotéz. Posléze implementace analýzy dat a vyhodnocení ve statistickém SW. Práce obsahuje 52 stran textu s obrázky a tabulkami. Práce je rozdělena do tří celků. V prvním diplomant popisuje technologie používané v rámci projektu MRE, v druhé statistické metody a ve třetí popis prováděné analýzy dat. V první i druhé části jsou popisy z větší části velice stručné, nedostačující a zvláště ve statistické části bych očekával alespoň náznaky toho, k čemu se daná metoda používá, jaké má vlastnosti, kdy ji můžeme/nemůžeme použít a proč. Většina popisů statistických metod se zakládá na vzorci a jeho krátkém slovním popisu (např.  $A = B + C$  – metoda sečte čísla B a C). Ve většině se přitom nejedná o běžně známé a používané pojmy. Stejný problém má i větší část kapitoly 4 – Analýza dat. I po několikátém přečtení mi stále nebylo jasné co vlastně diplomant dělal a hlavně proč. V celé práci se vždy objeví nějaký test, občas jeho popis a výsledek, jímž je nic neříkající graf v mizerné velikosti. Část grafů je navíc buď nesmyslná, nebo má otočené osy (diplomant v grafech v celé práci tvrdí, že spousta lidí má tlak přes 500, atp.).

Realizační část diplomant pojal jako generátor skriptů v jazyce java/bash pro statistické nástroje Matlab a R. Osobně nechápu, proč se k tomuto řešení diplomat uchýlil, když by šlo vše naprogramovat v jednom jazyce. Dalo by se také říci, že toto je nejslabší část celé práce, protože zdrojové kódy jsou nepřehledné, neokomentované a odporují jakýmkoliv pravidlům softwarového inženýrství a jejich úroveň je na úrovni prvního ročníku střední školy (žádné metody, celý program je psán v jedné třídě a metodě main, místo cyklů opakování bloků, špatné použití if, stejné bloky kódu se opakují místo volání metody, atp.). V práci zmíněnou interaktivitu konzolové aplikace jsem také marně hledal. Skripty v BASHi tento standart jen potvrzují, kdy akce, které můžeme provést jedním/dvěma příkazy diplomant provádí voláním několika programů. Výsledky to však vrací správně. Rovněž mi chybí jakákoliv dokumentace k použití programů.

## **Formální úroveň:**

Po formální stránce je práce v pořádku. Téměř neobsahuje překlepy ani gramatické chyby. Formátování práce je v pořádku. Výtku bych měl k publikovaným obrázkům, které jsou dost nepřehledné a nekvalitní.

## **Práce s literaturou:**

Práce obsahuje dostatečné množství správně citované literatury. Pouze bych předpokládal u dané problematiky alespoň nějaké tištěné zdroje.

**SOUHLASÍ  
S ORIGINÁLEM**



**Splnění zadání:**

Body zadání 1 – 3 zadání byly splněny. Bod 4 – tedy Zhodnocení výsledků a porovnání výsledků s řešením jiných autorů pouze částečně. Je však uveden objektivní a reálný důvod ohledně porovnání výsledků. Bohužel v této oblasti není možné provádět objektivní srovnání výsledků při neexistenci stejného vstupního datasetu.

**Doplňující informace k práci:**

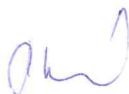
I přes výše zmíněné nedostatky, práci **doporučuji k obhajobě** a hodnotím klasifikačním stupněm slabší **dobře**.

V Plzni 5. 9. 2013



Ing. Michal Kratochvíl

**SOUHLASÍ  
S ORIGINÁLEM**



Západočeská univerzita v Plzni  
Fakulta aplikovaných věd  
katedra informatiky a výpočetní techniky  
②