

# Posudek oponenta bakalářské práce

Autor práce: **Vojtěch Jelínek**

Název práce: **Rozšíření možností Czech Salivary Gland Database pro klinickou praxi a pro analýzu dat**

Pan Vojtěch Jelínek se ve své bakalářské práci věnuje tématice shromažďování a zpracování klinických dat pacientů s diagnostikovaným nádorem slinných žláz ve spolupráci s Klinikou ORL a chirurgie hlavy a krku 1. LF UK a FN v Motole (dále jen ORL). Bakalářská práce úzce navazuje na semestrální práci, v rámci které byl vytvořen základ systému Czech Salivary Gland Database (dále jen CSGDB). Autor bakalářské práce je jedním ze tří spoluautorů této semestrální práce.

Cílem bakalářské práce je revidovat dosavadní funkcionalitu CSGDB ve spolupráci s uživateli na ORL, provést potřebné opravy, doplnění a rozšíření aplikace o výpočty a vizualizaci tzv. křivek přežití (Kaplan-Meierova metoda) s možností výběrů, resp. porovnání skupin pacientů.

Předložená bakalářská práce obsahuje celkem 87 stran, vč. příloh, rozsah vlastního textu je 65 stran (5 - 69), práce je členěna do 9 kapitol a dvou příloh, seznam literatury odkazuje na 15 zdrojů.

V první tematické části, kapitola 2, 3 a 4, se autor práce zabývá popisem původní verze CSGDB a analýzou jejích nedostatků. Protože tato verze aplikace CSGDB vznikla jako výstup semestrální práce, očekával jsem, že se autor bakalářské práce odkáže na původní dokumentaci. Takový odkaz jsem však nenašel. Proto mi připadá popis problematiky, která je řešena v bakalářské práci, nedostatečný tím spíše, že autor „začíná“ popisem uživatelského rozhraní a pak následuje analýza, resp. popis částí zdrojového kódu. Postrádám detailnější rozbor řešené úlohy, popis struktury a významu zaznamenávaných a analyzovaných dat. Např. přehled typů slinných žláz a terapií nebo popis histopatologie, které se např. dále využívají pro filtrování pacientů. Náznak takové informace obsahuje Tabulka 2.1, ovšem bez dalšího detailnějšího rozboru. S tím také souvisí nedostatečné vysvětlení pojmu „studie“.

V Tabulce 2.4 je uvedena klasifikace nádorového onemocnění. Z kontextu předpokládám, že se uvedená klasifikace týká nádorů slinných žláz. Je škoda, že autor nepředstavil zdrojový kód funkce *calculateTNMClassification*, podle které určuje stádium onemocnění. Také by bylo dobré doplnit odkaz na aktuální verzi *TNM Classification of Malignant Tumours*.

Analýza původní verze CSGDB, kterou autor provádí v interakci s uživateli na ORL, je rozdělena z hlediska funkcionality (používá tři stupně závažnosti) a z hlediska nedostatků v implementaci a struktuře kódu. Detekované nedostatky jsou popsány a současně je vždy uveden způsob jejich řešení, doplnění nebo opravy. Za zdůraznění stojí funkcionalita *Anonymizovaný export a Zabezpečení aplikace*, které byly z důvodu absence v původní implementaci autorem zařazeny do vysokého stupně závažnosti nedostatků. Byly následně řešeny podle pravidel, která jsou dále uvedena v kapitole 4, Problematika tvorby zdravotnických databází. Tím, že se autor při řešení uvedených funkcionalit neodkazuje na 4. kapitolu, působí pak samotná 4. kapitola v textu práce samoučelně, i když je obecně velmi významná a v jejím samotném závěru nakonec najdeme vysvětlení jejího smyslu vzhledem k aplikaci CSGDB.

Jako druhou tematickou oblast považuji implementaci výpočtu a vizualizaci tzv. metody Kaplan-Meier pro neparametrický odhad funkce přežití a test shody funkce přežití (Cox-Mantel, log rank test), kapitola 5. Zde ve vzorci (5.2) postrádám vysvětlení významu parametru  $T$  a ve vzorci (5.3) pak význam parametru  $R_i$ . V Tabulce 5.1, sloupec *Pravděpodobnost úmrtí*, by bylo pro úplnost dobré uvést vztah pro výpočet pravděpodobnosti úmrtí, známe-li pravděpodobnost přežití.

V případě Log rank testu jsou uvedeny vztahy pro dvě skupiny, dva výběry. Existuje ale také zobecněný log rank test pro porovnání více skupin. Jaká verze je implementována v aplikaci CSGDB? Co v této kapitole podle mého soudu zcela chybí, je graf odhadu funkce přežití, ať už ilustrativní nebo podle skutečných anonymizovaných dat pacientů ORL s náležitým vysvětlujícím komentářem (z textu bakalářské práce v kapitole 5 není jasné, zda bude grafické zobrazení teprve vytvářeno nebo již existuje, viz str. 35; na str. 58, Obrázek 6.9, je však ilustrativní křivka přežití zobrazena)!

Jako třetí téma bakalářské práce chápu implementaci kódu a nasazení aplikace, tj. kapitoly 6 a 7. Zejména kapitola 6, Implementace kódu, je velmi podrobná a rozsáhlá, která si vyžádala hodně systematické práce, což velmi chválím. Autor rozdělil implementaci celkem do sedmi postupných (iteračních) kroků, kdy každé rozšíření aplikace CSGDB podrobuje testování a oponentuře uživateli na

ORL, vypisuje nedostatky a chyby, které operativně řeší v definovaných časových úsecích. Výsledky celkového testování jsou shrnuty v Tabulce 6.9 a na Obrázku 6.10.

Příloha – A, str. 71-75, je uživatelským manuálem aplikace CSGDB, příloha – B, str. 77-78, popisuje obsah souboru A21B0154P\_prilohy.zip.

Po formální stránce je práce na dobré úrovni, i když mám výhrady ke struktuře a reálnému obsahu kapitol v první tematické části, viz poznámky v textu posudku. Dále práce obsahuje celou řadu překlepů a výrazných gramatických „zvláštností“.

Využití literatury je v přiměřeném rozsahu, všechny odkazy jsou funkční, kromě [GKK10]. Tato publikace je však snadno dohledatelná.

I přes výše uvedené připomínky považuji zadání bakalářské práce za splněné. Představená verze aplikace CSGDB je funkční.

Práci považuji za velmi přínosnou a vhodnou pro další rozvíjení, viz kapitola 8, Zhodnocení výsledků. Tomu také nasvědčuje předpokládaná účast na kongresu „10., 85., 70. ORL 2024“.

Vzhledem k výše uvedeným nedostatkům, které však považuji za méně závažné, navrhuji hodnocení známkou **velmi dobře**. Práci **doporučuji** k obhajobě.

Předpokládám, že pan Vojtěch Jelínek nejasnosti a dotazy uváděné v tomto posudku detailně během prezentace práce a její obhajoby vysvětlí.

V Plzni 24. 5. 2024

Ing. Pavel Nový, Ph.D.

*oponent bakalářské práce*

Katedra informatiky a výpočetní techniky

Fakulta aplikovaných věd

Technická 8

301 00 Plzeň

tel.: 377 632 411

e-mail: [novyp@kiv.zcu.cz](mailto:novyp@kiv.zcu.cz)