



České vysoké učení technické v Praze, Fakulta biomedicínského inženýrství
Katedra biomedicínské techniky, nám. Sítná 3105, 272 01 Kladno
tel.: +420 224 359 901, www.fbmi.cvut.cz
e-mail: nikola.lukacova@fbmi.cvut.cz

Studijní program „Biomedicínská a klinická technika“
studijní obor „Biomedicínský technik“

OPONENTSKÝ POSUDEK BAKALÁŘSKÉ PRÁCE

studenta: Anzhelika Kamkina

s názvem: Metody pro klasifikaci dat pacientů s podezřením na nádor žaludku

Hodnocení bakalářské práce dosahuje následující úrovně:

	Kritéria hodnocení bakalářské práce	Počet bodů
1.	Splnění cíle a vhodnost struktury obsahu bakalářské práce z hlediska zadaného tématu (splnění zadání). (0 - 30)* Každá část či věta ze zadání musí mít jasný odraz ve zpracované práci. Excelentně splněné zadání může být ohodnoceno maximálním počtem bodů. V poměru rozsahu části v zadání, která není zcela vhodně či úplně zpracována, se hodnocení odpovídajícím způsobem snižuje.	29
2.	Teoretická úroveň a využití dostupné literatury v bakalářské práci. (0 - 30) Oponent posuzuje relevantnost teoretické části k zadání, rozsah rešerší a systematické uspořádání zjištěných poznatků. Pokud převažuje doslovné převzetí textů, snižuje oponent hodnocení až o 15 bodů (přirozeně za předpokladu dodržení autorských práv). Důvodem pro snížení celkového hodnocení je dále nedostatečný výběr teoretických poznatků, literatury a zdrojů.	28
3.	Rozsah realizačních prací (SW, HW), aplikovaných vědomostí a znalostí, úroveň metodologického zpracování a závěrů práce. (0 - 30) Maximální počet bodů lze udělit práci, která je vhodná k publikování. Tento aspekt se posuzuje zejména z hlediska významu pro obohacení teoretických poznatků a má praktický význam. Obzvláště pozitivně je hodnoceno vytvoření modelu, SW produktu a též technická realizace. Za drobné metodologické nedostatky se hodnocení snižuje až o 5 bodů. Nekonzistentnost zpracování s teoretickými východisky a nejasný či ne zcela odborný metodologický přístup vede ke snížení minimálně o 15 bodů. Další snížení hodnocení lze udělit za nedostatečnou diskusi k závěrům. Celkem 30 bodů za velmi komplexní a bezchybnou práci včetně dalších aktivit jako je účast na vědecko-výzkumném projektu či grantu, aktivní účast na tvorbě publikací, patentů či užžitých vzorů.	28
4.	Formální náležitosti a úprava bakalářské práce (úroveň psaní, označení struktury textu, grafy, tabulky, citace v textu, seznam použité literatury apod.). (0 - 10) Oponent hodnotí formální náležitosti z pohledu dodržení pravidel o psaní, atributů závěrečných prací, tj. formátování textu, struktury práce, seznamu použité literatury, vybavenosti bakalářské práce grafy a tabulkami, způsobu citování. Za nedodržení jednotlivých pravidel snižuje maximální hodnocení o 2 body za každý nerespektovaný atribut. Rovněž za výskyt gramatických chyb, překlepů a nevhodné stylistiky a terminologie se snižuje hodnocení o 2-4 body. V práci by se měla objevovat pouze standardní odborná terminologie a to zejména v českém jazyce (je třeba hodnotit schopnost vyjadřovat se technickým jazykem - 2 body), grafy jsou tvořeny podle zásad (viz tolerance a vliv statistického zpracování - 2 body), u grafů a tabulek jsou patřičné legendy a vše je čitelné (2 body), jsou dodržena citační pravidla podle ISO690 a ISO690-2 (2 bod).	8
5.	Celkový počet bodů	93

Návrh otázek k obhajobě

1. Jako nejlepší se ukázala metoda rozhodovacího stromu spolu s metodou selekce ReliefF. Ale i rozhodovací strom vlastně provádí selekci příznaků. Jak spolu vybrané příznaky souvisely?
2. Kterým atributům přiřadila nejlepší neuronová síť nejvyšší (absolutní hodnotu) vah v první vrstvě? Jak je to v porovnání s příznaky, vybranými jinými metodami selekce?
3. Jak se lékaři dívají na získané výsledky? Zajímá je vyšší sensitivita nebo specifita testu? Jak komentovali nejdůležitější vybrané atributy, např. atributy nejlepšího rozhodovacího stromu? Jaký je význam nejlepšího získaného rozhodovacího stromu z klinického hlediska?

Celkové hodnocení úrovně vypracování bakalářské práce:

Hodnocení**:	A (výborně)	B (velmi dobře)	C (dobře)	D (uspokojivě)	E (dostatečně)	F (nedostatečně)
Počet bodů:	100 - 90	89 - 80	79 - 70	69 - 60	59 - 50	< 50
	X	□	□	□	□	□

** v případě hodnocení F (nedostatečně) uveďte komentář

Bakalářskou práci hodnotím výše uvedeným klasifikačním stupněm a doporučuji/~~nedoporučuji~~ k obhajobě.

Komentář

Po stránce machine-learningu je práce naprosto v pořádku, v některých ohledech se vyrovná i některým diplomovým pracím (metodologie a množství odvedené práce).

V práci postrádám úvodní analýzu atributů, např. histogramy a scatter-plot. Pokud je uveden průměr, je vhodné uvést i směrodatnou odchylku (týká se mnoha tabulek). Nicméně u těch důležitějších tabulek uvedena je.

V Kapitole 3.1 (anebo alespoň kap. 5) chybí popis atributů z (bio)medicínského hlediska, není jasné, které jsou klinicky významné. Z práce mám občas dojem, že je vše bráno pouze z pohledu data-minera, ale chybí zpětná (a možná i dopředná) vazba experta-lékaře.

U obr. 15 (rozhodovací strom) chybí kvantitativní ohodnocení uzlů (počet instancí (dle tříd), které se dostanou do jednotlivého uzlu). Navíc, hodnoty by bylo vhodné denormalizovat, pokud mají být předány lékaři: např. prahová hodnota věku je v práci: 0.098.

Největším nedostatkem je malý počet instancí dat (vzhledem k počtu atributů), ale to nejspíš není chyba autora práce.

Jméno a příjmení: Ing. Miroslav Burša, Ph.D.
Organizace: Český institut informatiky, robotiky a kybernetiky
Kontaktní adresa: Evropská 11, 160 00 Praha 6

Podpis:
Datum: