



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: Barbora Balcarová

s názvem: Analýza příznaků automatické klasifikace epileptických EEG záznamů za pomoci algoritmu k-NN

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.	30
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.	30
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 užitečných vzorů.	20
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).	10
5.	Celkový počet bodů	90

Návrh otázek k obhajobě

1. Normalizace popsaná na straně 19 není dle mého názoru úplně korektní. Tvrzení: "Pro normalizaci je však třeba použít celý pacientův EEG záznam," neodpovídá realitě. Dle mého názoru je třeba vždy počítat normalizační transformaci pouze z trénovacích dat a tuto transformaci použít pro normalizaci všech, tedy jak trénovacích tak testovacích dat. Důvod je ten, že při reálném použití neznáme testovací data dopředu a pro normalizaci je tedy využít nelze. Jedná se však o malou chybu, která však může optimisticky vychýlit výsledný odhad chyby klasifikátoru. Okomentujte.

2. Podobně je tomu u tvrzení: "Korelační koeficienty jsem počítala pro každou dvojici příznaků celého patientského záznamu." Toto je metodologicky ne zcela správné, protože testovací množina by zde pro výpočet korelačních koeficientů neměla být použita. V reálu selektujeme příznaky pouze za pomoci dostupných dat, kterými nejsou data, která získáme v průběhu užívání klasifikátoru. Nesprávně předpokládaná znalost těchto "budoucích dat" může způsobit optimistické vychýlení odhadu chyby. Opět okomentujte.

3. Podle 4. a 5. řádku od konce Algoritmus 1 nepopisuje pouze kNN, ale jeho zvláštní verzi, kdy klasifikovaná data jsou přidávána do trénovací množiny. Jedná se tak o jakousi verzi semi-supervised learning (viz. https://en.wikipedia.org/wiki/Semi-supervised_learning). Jak se liší původní algoritmus kNN od algoritmu 1?

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	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

** v případě hodnocení F (nedostatečně) uveďte komentář

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

Komentář

Práce je zaměřena na klasifikaci EEG segmentů do dvou tříd pomocí klasifikátoru 5NN. Popisuje analýzu semiautomatické metody, kdy veškerá data pro trénování i testování automatického klasifikátoru jsou předem ručně předvybrána a nazvána etalony. To dle mého názoru trochu snižuje využitelnost výsledku, protože v reálu máme možnost takto předvybrat pouze trénovací data. Etalony by měly tvořit pouze trénovací množinu, testovací množina by měla být vybrána náhodně. Hlavním bodem analýzy je sledování vlivu eliminace vzájemně korelovaných příznaků.

Práce je po formální stránce velmi pěkná. Metodologicky se zde vyskytují menší nedostatky. Výsledky jsou pěkně shrnuty a diskutovány.

Jméno a příjmení: Ing. Martin Macaš, Ph.D.

Organizace: ČVUT v Praze, CIIRC

Kontaktní adresa: Jugoslávských partyzánů 1580/3, 160 00 Dejvice

Podpis:

Datum: