

Posudek oponenta diplomové práce

Název práce: Zpracování videookulografických dat se zaměřením na detekci dyslexie

Posudek vypracoval oponent práce: Ing. Martin MACAŠ, Ph.D.

Český institut informatiky, robotiky a kybernetiky, ČVUT

Autorka práce: Bc. Alena Juklová

Práce je trochu nevyváženě strukturována do pěti kapitol. Kapitola 2 řeší první bod zadání práce provedením rešerše literatury týkající se diagnostiky dyslexie. Autorka se věnuje hlavně popisu oka, očních pohybů a jejich měřením. Dále jsou shrnuty obecné metody statistického zpracování dat. Ačkoliv jsou zmíněny a citovány některé spojitosti mezi dyslexií a očními pohyby, rešerše by zde měla být mnohem důkladnější a soustředit se také na nejnovější publikace z posledních několika let. Měl by také být kladen větší důraz na onu diagnostiku očních pohybů a její obtíže. Na straně 22 je v popisu křížové validace chybně uvedeno, že „trénovací množina je náhodně rozdělena do m disjunktních podmnožin stejné velikosti m/n , kde n je počet vzorků“.

Kapitola 3 nejen popisuje použité metody, ale uvádí také výsledky statistického porovnání příznaků pro skupiny dyslektiků a ne-dyslektiků. Zde oceňuji ověření předpokladů analýzy rozptylu. Analýza rozptylu samotná pak ukazuje zajímavé signifikantní rozdíly pro některé příznaky. Kapitola 3.6.3 popisuje použití jednoduchého klasifikátoru s jedním vstupem. Klasifikátor spočítá průměrné hodnoty příznaku pro každou třídu a rozhodovací práh nastaví na hodnotu průměru těchto dvou hodnot. Při popisu tohoto klasifikátoru na straně 41 je nesprávně uvedeno, že se jedná o učení bez učitele, protože informace o třídách je při nastavování prahu použita. Dále zde není jasné, jakého odhadu přesnosti, sensitivity a specificity bylo použito, tedy **na jakých datových množinách byl klasifikátor učen a na jakých byl testován. Toto by měla autorka osvětlit během obhajoby**. Kapitola 3.6.3 pak popisuje složitější vícerozměrové klasifikátory s několika vstupními příznaky, které jsou automaticky vybrány pomocí tzv. wrapper metody. Klasifikátory jsou vyjmenované ale chybí detailnější popis nebo reference. Zde naopak chválím detailní popis testování klasifikátoru pomocí leave-one-out metody. Výsledky klasifikace jsou sumarizovány a pěkně srovnány s předchozími výsledky statistického testování. Kapitola 4 je nevyváženě krátká a mohla by být včleněna do kapitoly 3. Autorkou zmiňovaný problém s nízkou sensitivitou je patrně způsoben nevyváženou reprezentací jednotlivých tříd v datové množině a může být redukován laděním klasifikátorů za cenu snížení specificity. Vysoká výkonnost jednoduchého lineárního Bayesovského klasifikátoru je zřejmě způsobena malým datovým vzorkem. Závěr práce uvedený v kapitole 5 pak shrnuje všechny výsledky.

Po formální stránce mám jen drobné výhrady. Některé obrázky (např. 3.4 a 3.5 jsou velmi nekvalitní (špatné rozlišení, nečitelné texty v obrázcích). Citace jsou někdy nesprávně uváděny až za tečkou ukončující větu.

Všechny body zadání byly splněny. Větší výhrady mám jen k bodu prvnímu – rešerši, která dle mého názoru mohla být důkladnější, aktuálnější a relevantnější. Přes všechny uvedené výhrady se mi práce velmi líbí, autorka k řešení i popisu přistoupila poctivě a dokázala, že je schopna kvalitní a samostatné inženýrské práce. Předloženou diplomovou práci proto hodnotím stupněm

B - velmi dobře