

“Klastrování ručně psaných znaků”

Předkládaná diplomová práce se zabývá problematikou automatického rozpoznávání ručně psaných znaků. Z pohledu klasické teorie rozpoznávání (obrazových 2D) vzorů patří tato problematika k víceméně vyřešeným. V praxi běžně používané systémy pro rozpoznávání ručně psaných znaků dosahují přesnosti nad 90% (jinak by pro zpracování ručně psaných dokumentů nebyly vůbec použitelné) a další zvýšení jejich spolehlivosti lze očekávat především ve spojení se statistickými metodami pro zpracování přirozeného jazyka, resp. se strukturální analýzou rozpoznávaných znaků.

Jako další vhodná alternativa připadají v úvahu i novější modely hybridních anebo konvolučních neuronových založených na masivně paralelním zpracování předkládaných vzorů, které však nevyžaduje náročné předzpracování vstupních dat. Některé z těchto modelů využívají i principy klastrování.

Ve své diplomové práci se měl posluchač zaměřit na oblast rozpoznávání ručně psaných znaků pomocí metod založených na principu klastrování. Posluchač měl nastudovat standardní klastrovací techniky (např. sekvenční metody) i modely neuronových sítí (např. Kohonenovy mapy, hybridní model neuronové sítě), které by bylo možné použít pro rozpoznávání ručně psaných znaků.

Vybrané techniky a modely měl implementovat a otestovat (např. na ručně psaných číslicích). Součástí práce mělo být jednak experimentální ověření a porovnání vybraných modelů na dané problematice i srovnání jednotlivých klastrovacích technik s ohledem na různé parametry technik (např. různé kriteriální funkce). Práce měla obsahovat i zhodnocení vlastních výsledků a zkušeností.

Předkládaná práce je velmi rozsáhlá, z věcného hlediska má však celou řadu vážných nedostatků. Diplomant se pokusil nastínit základní principy některých ze standardních klastrovacích technik i klasických modelů umělých neuronových sítí. Popis těchto metod však obsahuje celou řadu chyb. Mnoho z nich je neslučitelných se správnou funkcí příslušných modelů – viz např.:

- předpoklad normálního rozdělení číslic na str. 38: podle uvedeného popisu měl autor patrně na mysli rovnoměrné rozdělení; daný problém mu ale neodpovídá, protože psaná čísla nejspíš nebudou začínat '0',
- chyby v pseudokódu pro Algoritmus k nejbližších sousedů: cyklus by měl probíhat přes seřazené reprezentanty; implementovaná technika je navíc neefektivní, i když pro autorem napevno zvolenou hodnotu $k=2$ se tento efekt při testování neprojeví; dále viz str. 41,
- chyby v popisu Algoritmu zpětného šíření chyby: ve vzorci 5.2 je třeba nastavit $w_0=h$; hodnoty prahů je třeba aktualizovat i pro výstupní neurony; ve vztahu 5.14 chybí parametr učení; dále viz str. 55-56,
- pevně nastavený počet shluků odpovídající počtu požadovaných tříd v Kapitole 6: je vhodnější použít větší počet shluků a vytvořené shluky následně okalibrovat; stejnou

třídou je možné okalibrovat i vícero shluků, pokud reprezentují různé formy téhož znaku,

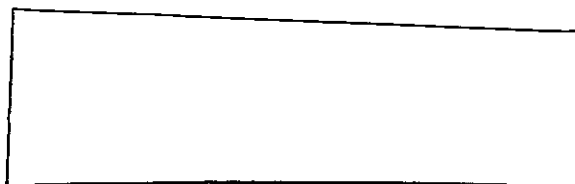
- analýza modelu hybridní neuronové sítě, který byl uvedený v zadání, v práci chybí úplně.

Techniky, které uchazeč ve své práci nakonec pro rozpoznávání ručně psaných znaků zvolil, pro řešení zvolené úlohy použitelné nejsou, a to včetně metod zvolených pro předzpracování vstupních dat a extrakci příznaků. To ostatně potvrzují i výsledky provedených experimentů. Kromě uváděné 39%-ní přesnosti klasifikace v této souvislosti budí velké rozpaky i tabulky výsledků klasifikace uváděné na různých místech v Kapitole 6 (např. na str. 70) a „ruční zhodnocení výsledků“ zmiňované v posledním odstavci na str. 87. Průběh experimentálního ověřování vlastností vybraných modelů s ohledem na různé parametry technik patrně naznačují zakomentované bloky kódu na přiloženém CD – viz soubor Nsit.cpp. Pro testování by bylo navíc vhodnější použít reprezentativnější data, např. databázi MNIST dostupnou na <http://yann.lecun.com/exdb/mnist/>.

I po formální stránce má předkládaná práce mnoho nedostatků. Často se v ní objevují překlepy a gramatické chyby – jen na str. 88 jich je 10, dále viz např. ř. 2. sh. na str. 38, ř. 4 zd. na str. 47, ř. 7 zd. na str. 61, ř. 9 a 11 zd. na str. 87 – text však zůstává víceméně srozumitelný. Výhrady lze mít i ke grafické úpravě textu – viz např. tabulky přesahující do okrajů stránek (6.1 na str. 67, 6.13 a 6.14 na str. 83), resp. jediné slovo zbývající na str. 89. Seznam literatury na str. 92 obsahuje neúplné reference, některé z pramenů ani nejsou dostupné.

Závěrem, bohužel, musím konstatovat, že předkládaná práce nesplňuje původně stanovené cíle. Protože práce Jiřího Nováka nesplňuje požadavky kladené na diplomovou práci, nedoporučuji uznat ji jako práci diplomovou.

V Praze, 26. 1. 2009



Doc. RNDr. Iveta Mrázová, CSc.
KTIML MFF UK