

Vyjádření vedoucího k diplomové práci

Petra Tobolíková: Detekce chyb v rozpoznávání mluvené řeči

P. Tobolíková se ve své práci zbývá aktuálním tématem aposteriorní detekce chyb ve výstupu rozpoznávače mluvené řeči. Tato tematika se doplňuje se snahou o zlepšení rozpoznávačů obecně, a to z pozice přehodnocení výsledků rozpoznávače při znalosti jeho některých vnitřních parametrů.

Práce obsahuje úvod (kap. 1), ve kterém autorka nastiňuje základní problematiku rozpoznávání mluvené řeči. Ve druhé kapitole pak podává přehled současných metod rozpoznávání, a zejména se věnuje metodám založeným na skrytých Markovových modelech (tj. metodám v současnosti nejvíce používaným) a jejich parametrizaci a trénování. Ve třetí kapitole popisuje metody aposteriorní detekce chyb na výstupu rozpoznávačů, které rozděluje na pravděpodobnostní a ostatní. Ve čtvrté kapitole pak popisuje možnosti aplikace strojového učení na problém detekce chyb rozpoznávačů. V páté kapitole popisuje vlastní experimenty a implementaci detekce chyb, pro kterou využívá tři různé metody: logistickou regresi, neuronové sítě a rozhodovací stromy, a rovněž využívá i jejich kombinace. V této kapitole také dosažené experimenty vyhodnocuje formou tabulek, porovnání (možné) dosažené chybovosti pomocí standardu Word Error Rate a porovnává tyto výsledky s jinými publikovanými metodami. Šestá kapitola je pak dokumentací její implementace (uživatelský popis rozhraní a dokumentace ke kódu). V závěru práce (kap. 7) pak autorka shrnuje dosažené výsledky, uvádí seznam nejvhodnějších rysů (kontextu) pro detekci chyb, a naznačuje i další možnosti, jak detekci chyb dále zlepšovat; konstatuje, že na základě jejích experimentů se jako nejvhodnější (samostatná) metoda detekce jeví metoda rozhodovacích stromů. V příloze uvádí příklad výstupu rozpoznávače a tabulku směrodatných odchylek v provedených experimentech. K práci je přiloženo CD s programy, daty a dokumentací.

Hodnocení:

Práce je psána česky, čte se velmi dobře jak v úvodních částech, kde autorka popisuje současný stav poznání v oblasti rozpoznávání mluvené řeči, tak v částí, kde popisuje vlastní přínos k této problematice. Při prováděných experimentech postupovala v souladu s uznávanými standardy (Word Error Rate jako míra chyby, způsob získávání parametrů v metodách strojového učení, vyhodnocení výsledků, stanovení směrodatných odchylek). Pro vyhodnocování výsledků vlastní detekce chyb používala inovativně a jednoduše tzv. Receiver Operating Characteristics (křivka ROC), která je vhodná na daný typ detekce a umožňuje porovnání jednotlivých metod a experimentů v celém rozsahu poměru skutečných a falešných pozitiv. Je třeba kladně hodnotit i velké množství experimentů, které musela autorka provést, aby dostala dostatečné množství výsledků k porovnání a výběru nejlepší metody.

Závěr: celkově práci považuji za vynikající a doporučuji, aby byla přijata a obhájena jako práce diplomová.



Praha, 10. 5. 2007, Jan Hajič, UFAL MFF UK