

Oponentský posudek disertační práce

Ing. Tomáš Mikolov: Statistical Language Models Based On Neural Networks

Práce Tomáše Mikolova je zacílena na jazykové modelování, kde po dlouhou dobu zůstávaly nejúspěšnější tzv. N-gramové modely, přes jejich zřejmou neadekvátnost z jazykového hlediska. Jeho cílem bylo najít nové modely a takové metody jejich použití, které by zůstávaly i z technologického hlediska použitelné v aplikacích (kde např. rozpoznávání řeči vyžaduje téměř vždy práci v reálném čase). Pro prověření těchto vlastností byly nově vytvořené modely evaluovány nejen teoretickými mírami, ale i v několika standardních aplikacích (ASR) a na standardních, tj. srovnatelných souborech evaluačních dat (např. Penn Treebank).

Práce má 133 stran a je rozdělena do devíti kapitol. Po nezbytném úvodu pokračuje úvodem do statistického jazykového modelování, které netradičně, ale zcela správně začíná popisem evaluačních metod (jak ostatně vždy zdůrazňoval F. Jelinek, jeden z pionýrů použití statistických metod ve zpracování řeči a zpracování jazyka a zejména použití objektivních evaluačních metod). Třetí kapitola se zaměřuje na zvolenou třídu jazykových modelů, tj. na jazykové modely založené na neuronových sítích. Velká část této kapitoly je věnována rozšíření těchto metod, které bylo v jádru práce autora. V závěru této kapitoly jsou i podrobně popsána testovací (evaluační) data použitá v této práci. Čtvrtá kapitola popisuje experimenty provedené autorem, včetně výsledků na standardních datových souborech. Pátá kapitola pak popisuje aplikace na úlohu rozpoznávání mluvené řeči, a jsou zde rovněž popsány výsledky na evaluačních souborech dat (opět vybraných standardně tak, aby bylo možné přímé srovnání s předchozími výsledky). V šesté kapitole pak autor opisuje rozšíření jazykových modelů založených na neuronových sítích z hlediska efektivního trénování těchto modelů z velmi velkých dat (až 400 mil. slov) a výsledky opět prezentuje na standardních evaluačních datech (NIST data sets). V sedmé kapitole jsou popsány další experimenty s aplikacemi, kde se také typicky používají jazykové modely – strojový překlad, komprese data apod. V osmé kapitole se autor zamýšlí nad omezeními jazykových modelů obecně a navrhuje některá řešení do budoucna. Poslední, devátá kapitola, pak shrnuje dosažené výsledky. Práce je doplněna přílohami (open source toolkit – dokumentace, příklady z Broadcast News dat).

Hodnocení:

Předložená práce představuje posun v oblasti jazykového modelování, a to v absolutním měřítku (tj. ve světovém srovnání). Autorovi se navíc podařilo dosáhnout výjimečné symbiózy nového postupu (neuronové sítě a jejich rozšíření, které jsou v této kombinaci nové řešení, i když neuronové sítě jako takové se používaly pro jazykové modelování již dříve, jak ostatně autor řádně cituje) a významně zlepšených výsledků, a to i aplikačních. Podařilo se mu i přesvědčivě ukázat, že neuronové sítě, považované často za tak velmi výpočetně náročné (až nepoužitelné) v praktických aplikacích, mohou být pro jazykové modelování použity efektivně. Významným výsledkem práce je i poskytnutí vytvořeného software jako open source, tj. na práci bude možné hladce navázat v dalším výzkumu. Práce je psána velmi dobrou angličtinou a ani z formálního hlediska jí není co vytknout.

Závěr:

Bez jakýchkoliv pochybností doporučuji, aby byl komisí pro obhajoby doktorských prací na FIT VUT Brno Ing. Tomáši Mikolovi udělen titul Ph.D.



Praha, 17. 9. 2012, Jan Hajič, ÚFAL MFF UK